# Effective Network Quarantine with Minimal Restrictions on Communication Activities

Huanyang Zheng and Jie Wu, *Fellow, IEEE*

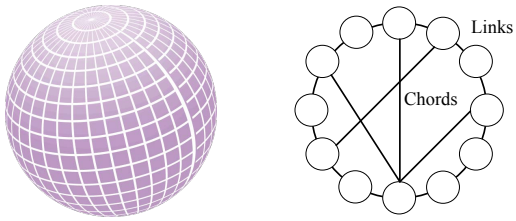
**Abstract**—This paper studies a network structure design problem constrained by the epidemic outbreaks. In our model, geographic locations (nodes) and their connections (edges) are modeled as a ring graph. The movement of a person is represented as a flow from one location to another. A person can be infected at a location (node), depending on the number of infected flows (persons) going through that location. In our paper, diseases are not limited to real human diseases; they can also refer to the general epidemic information propagations. Given desired interaction traffic from a node to other nodes in terms of flows, and a greedy shortest path routing scheme that is analogous to the greedy coin change, we focus on the structure design (representing quarantine rules) that determines the number and distribution of chords on the virtual ring network for remote connections. Our objective is to minimize the average number of routing hops, while the epidemic outbreaks are controlled under given infection and recovery rates. We provide a systematic isomorphic structure design on nine different cases, based on three traffic distribution and three infection rate models. Two hypercube-based structures are proposed. We also provide a greedy solution for constructing polymorphic structures. Our study reveals some intriguing theoretical results, validated through experiments, on tradeoffs between local and global infections. Our work casts new light on the effective network quarantine that places minimal restrictions on connections, i.e., maximal preservation of normal communication activities, while controlling epidemic outbreaks.

**Index Terms**—Epidemic outbreaks, interaction, quarantine, network structure design.

## 1 Introduction

The notion of the disease [1] has been extended from human diseases, such as Ebola, to general epidemic information propagations [2], such as rumors in distributed networks. Controlling the spread of a communicable disease in a population is usually done through the *quarantine*, where persons that have, or are suspected to have, a communicable disease are restricted from having interactions with others. However, a quarantine is inconvenient for communication activities, due to its limitations on interactions. Hence, there is a tradeoff between the control of *epidemic outbreaks* and the convenience of communication activities. To understand this tradeoff, we conduct studies based on a virtual ring model, which is an abstraction of the real world through meridian and parallel. As shown in Fig. 1(a), meridian and parallel discretize the Earth into geographic locations (represented by nodes) on multiple rings. For each ring, neighboring nodes are connected via *links* (roads between nearby towns). Geographically-remote nodes are connected via *chords* (flights between remote cities). The geographical distance between nodes can be measured by their hop distances along the ring.

This paper studies a network structure design problem constrained by the epidemic outbreaks. In our model, flows represent the movements of people. A person can be infected at a location (node), depending on the number of infected flows (persons) going through that location. Person-to-person interactions are controlled by flow routes among different locations. The quarantine is represented as interactions with neighbors only (through a link in a 1-D ring, say, walking between two adjacent towns), and non-quarantine is shown as interactions with remote nodes (through a chord in a 1-D ring, say, flying between two cities). A real-world example is that Australia clamped down on the entry of individuals traveling from West Africa [3], in order to control Ebola. Such a quarantine rule is executed by the Australian Customs and Border Protection Service. The SIS model [4] is used to simulate the epidemic spreading, where flows have states of being *susceptible* or *infected*. Flows transfer their states through a cycle of being infected (based on a given infection rate) from susceptible, and going back to susceptible by recovery (based on a given recovery rate). We consider three infection rate models, where the infection probability is related to the infected flows at a node. The recovery rate is constant, as used in existing models [4].



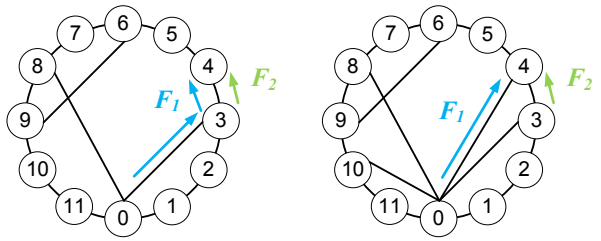

(a) Meridian and parallel. (b) Virtual ring model.

Fig. 1. Abstraction of the Earth with meridian and parallel.

We assume that each node has fixed outgoing traffic, in terms of flows, to other nodes, following a particular distribution [5]. Flow interactions occur in nodes along a routing path from source to destination, which is a sequence of links and chords. The routing path is determined by

• *H. Zheng and J. Wu are with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, 19122.*

*Manuscript received May 5, 2015; revised September 17, 2015.*




(a) Local infection. (b) Global infection.

Fig. 2. An example of the traffic flow routing.

a greedy-coin-change-based routing scheme [6], [7]. Each node will select a connected relay, which is the closest to the destination, to forward its traffic. In Fig. 2(a), if $N_0$ wants to interact with $N_5$ by generating a flow $F_1$, then its path is $\{N_0, N_3, N_4\}$. In general, each node has multiple flows, including *forwarding flows* and its own traffic flows. For example, $N_3$ has its own desired traffic to $N_4$ in the flow $F_2$, while forwarding the flow $F_1$ from $N_0$ to $N_4$.

Classic epidemiology controls the epidemic outbreaks by studying the space-time patterns of diseases [8]. In contrast, we study a network structure design problem constrained by the epidemic outbreaks. The designed network structure should *minimize the average number of routing hops for node interactions without epidemic outbreaks*. The application of our work implies a quarantine rule, which can maximally preserve normal communication activities without epidemic outbreaks. Note that the effect of a given flow is not restricted over its routing path. If we design a ring structure that has few chords, then the traffic will aggregate along local nodes. In this case, susceptible flows going through locations with infected flows are very likely to be infected. This phenomenon is called the *local infection*. For example, Fig. 2(a) has less chords than Fig. 2(b). As a result, $F_1$ and $F_2$ aggregate along the link from $N_3$ to $N_4$. Aggregated flows lead to a larger infection probability, and thus, local infection happens. On the other hand, if we design a ring structure that has numerous chords, then the traffic is spread widely through chords. An infected flow can reach many susceptible remote flows to spread the disease, leading to the *global infection* phenomenon. For example, infected flows from $N_0$ can directly reach more nodes in Fig. 2(b) than Fig. 2(a). Since flows interacts with epidemics, the tradeoff between local infection and global infection makes the structure design problem very challenging.

In this paper, we mainly focus on *isomorphic structures*, where each node on the ring has the same chord distribution in terms of degree and neighbor distribution. The sensitivity experiments demonstrate that the isomorphism assumption can be relaxed, and thus they are feasible for real-world applications. Table 1 shows our major findings, based on three traffic distribution models and three infection rate models. The traffic distribution models (indicating *different node communication patterns*) are shown as follows: (I) The traffic from one node to another node is a constant; (II) The traffic from one node to another node decays slowly with respect to their geographical distance; (III) The traffic from one node to another node decays quickly with respect to their geographical distance. The infection rate models (indicating *different epidemic infectivities*) include: (A) The probability of being infected by infected flows is a constant; (B) The probability of being infected is sub-linearly proportional to the traffic amount of infected flows; (C) The probability of being infected is linearly proportional to the traffic amount of infected flows. For a virtual ring with $n$ nodes, the properties in Table 1 are listed as follows:

TABLE 1
Major results for isomorphic structures

| Properties of (1-6) are itemized in the text. | | Traffic distribution models | | |
|---|---|---|---|---|
| | | Model I | Model II | Model III |
| Infection rate models | Model A | 1, 5 | 1 | 1, 4 |
| | Model B | 2 | 2, 6 | 2, 4 |
| | Model C | 3 | 3 | 3 |

1) The global infection is the major factor for the epidemic outbreak. Epidemic outbreaks could be controlled, if the node degree is capped.
2) Neither local nor global infection is the major factor for the epidemic outbreak. The optimal structure remains an open problem.
3) The local infection is the major factor for the epidemic outbreak. The fully-connected network is the optimal structure.
4) Chords should only connect to the geographically nearest $O(\ln n)$ nodes to facilitate the communications. This is because the interaction traffic decays exponentially with the geographical distance.
5) The Ulysses butterfly structure [9] is an asymptotically optimal structure.
6) This corresponds to the most complex case, and is the focus of this paper. Two hypercube-based structures are proposed and analyzed. The first structure uses binary jump sizes of chords [7]. The second structure improves the first one by considering the desired traffic distribution.

We also propose a greedy solution for constructing *polymorphic structures*, where nodes have heterogenous chord distributions. This greedy solution is studied under the desired traffic model I and the infection rate model A. We show the relationship between the reduced average number of routing hops (i.e., the benefit) and the increased network vulnerability to epidemics (i.e., the cost), brought by adding a chord. The chord, which has the largest benefit-to-cost ratio, is iteratively added to construct a polymorphic structure. This greedy solution gives out some insights on efficiently constructing general polymorphic structures that resist epidemic outbreaks.

The remainder of this paper is organized as follows. In Section 2, the basic model is described and the problem is formulated. In Section 3, two hypercube-based isomorphic structures are studied. In Section 4, we consider other traffic distribution and infection rate models. In Section 5, we focus on constructing polymorphic structures. In Section 6, our network structures are evaluated. In Section 7, we conclude the paper. Supplementary material are published in [10].

## 2 Motivation, Model, and Formulation

We first describe the motivation. Then, the ring network model, the traffic distribution models, and the infection rate models are described. Finally, the problem is formulated.

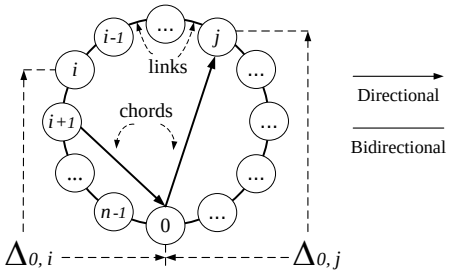



Fig. 3. The virtual ring model with bidirectional links.

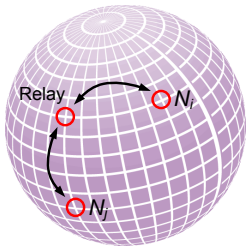



Fig. 4. A globe representing Earth with meridian and parallel.

### 2.1 Motivation

This paper is based on the real earth. Meridians and parallels discretize the earth into different locations (cities). People move around different locations for their lives. Epidemics (e.g., Ebola) are found and may spread. To avoid epidemic outbreak, people's movements are restricted by quarantine. For example, Australia clamped down on the entry of individuals traveling from West Africa to against Ebola [3]. However, some movements are necessary and desired, even under the threat of epidemics. For example, food suppliers have to move around to feed people. Consequently, it is challenging to find an answer to the question of how to control these necessary movements while avoiding epidemic outbreaks. Let us consider a motivational scenario on the transportation network for food suppliers:

- To feed people, food must be transported from some cities to some other cities by trains. These movements are necessary, even under the threat of epidemics.
- To avoid epidemic outbreaks, two cities may not be directly connected by trains. A transportation between them is forwarded via some third-party cities.

The structure of the transportation network (i.e., which two cities are directly connected by trains) for food suppliers is critical. It is the variable in this paper (represented by the number and distribution of chords). We want to satisfy the food supply requirement and control the epidemics, with minimal transports. If all cities are directly connected by trains, then an epidemic outbreak can spread globally. However, if only a few cities are directly connected by trains, then people need to go through multiple cities to gain access to food supplies. People who goes through more cities also have a higher possibility of being infected.

Cities on the earth are modeled by the ring network in subsection 2.3. Necessary people's movements (e.g., food supplies) are described by given traffic distribution models, with a routing scheme, in subsection 2.4. The traffic can be routed directly and indirectly (e.g., nonstop and multi-stop trains). This interacts with epidemic spreadings, which are described by given infection rate models in subsection 2.5. The problem is formulated in subsection 2.6. The objective is to minimize the number of routing hops for traffic flows, and the constraint is to control epidemic outbreaks.

### 2.2 Related Works

Our paper is closely related to transportation network design papers [5], [11], which design airline/train network structures to satisfy transportation demands and constraints [12], [13]. For example, Wieberneit [13] review the literature on designing the freight transportation network, including the network structure, the freight frequency, and the freight routing path. The freight service demands are satisfied. Our research builds upon previous knowledge, but additionally consider the constrains imposed by the epidemic outbreak. This paper aims at a network structure, which brings a minimum routing for given traffics, and the constraint is that epidemics will not outbreak. This paper is also related to peer-to-peer network structure design problems. For example, Xu's work [5] studies a degree-capped network structure that achieves a minimum diameter. Our problem becomes equivalent to Xu's work, when traffics are uniformly-distributed and the epidemic is traffic-insensitive. This is because the minimum routing becomes equivalent to the minimum diameter, and the epidemic constraint becomes equivalent to the degree cap.

The spreading of epidemics has been well-studied in the literature [14], [15], [16], [17], [18]. Classic models in epidemiology, such as SIS and SIR, are summarized in [4] with respect to different network structures. These models are known as compartmental models, which use interactions between states (e.g., susceptible, infected, recover) to describe epidemics. However, these models do not consider the interactions between the epidemic spreading and the traffic. Therefore, Preciado et al. [16], [17], [18], [19] introduced convex optimization techniques for the epidemic control under different traffic rates. In contrast, this paper uses the existing SIS model to study the interaction between the traffic flow routing path and the epidemic spreading through a network structure design problem.

### 2.3 The Ring Network Model

As shown in Fig. 3, our study is based on a virtual ring network of $n$ nodes. Since the virtual ring is an abstraction of a real-world geographic map, each node on the ring represents a geographic location. Nodes have links with two geographical neighbors, and chords with geographically-remote nodes. The IDs of nodes are from 0 to $n-1$, where the $i^{th}$ node is denoted as $N_i$. The geographical distance between $N_i$ and $N_j$ is denoted as $\triangle_{i,j}$, which is measured by their hop distances along the ring (the minor part). In Fig. 3, we have $\triangle_{0,j} = j$ and $\triangle_{0,i} = n-i$. Both links and chords are called edges in this paper, and are directional. We first explore isomorphic structures, where each node has a degree of $D$ (both in-degree and out-degree). Polymorphic structures are considered later in Section 5.

The ring model can represent the geographic relationships among different locations. It is widely acknowledged

in research on P2P networks [7] and social networks (the small-world phenomenon) [20]. As shown in Fig. 4, each meridian or parallel of the Earth can be abstracted as a ring. In such a case, node communications can be decomposed on two rings. In Fig. 4, $N_i$ and $N_j$ can communicate with each other through a relay node at the intersection of the meridian and the parallel rings (X-Y routing). More details are presented in the supplemental material.

Links in the ring represent the network backbone that maintains the network connectivity. Quarantine is represented as interactions with neighbors only (through a link in the ring, say, walking between two adjacent towns). People are always allowed to move near to their home locations. Chords stand for the normal connections. Non-quarantine is shown as the interactions with remote nodes (through a chord in the ring, say, flying between two cities). A real-world example is that Australia clamped down on the entry of individuals traveling from West Africa [3]. There is no chord from West Africa to Australia, i.e., people in West Africa are not allowed to travel to remote locations directly.

### 2.4 The Desired Traffic Distribution Model

In this subsection, we start with the desired traffic distribution model II. Let $T_{i,j}$ denote the desired interaction traffic from $N_i$ to $N_j$. $T_{i,j}$ can be interpreted as a flow, which includes a certain number of people moving from locations $N_i$ to $N_j$. $T_l$ is the total interaction traffic going through $N_l$ (including the traffic that heads for $N_l$, and the traffic that is forwarded by $N_l$). The mean value of $T_l$ is denoted as $\langle T_l \rangle$. Since people are more likely to move around their home locations, we consider $T_{i,j}$ to be $T_{i,j} \propto \frac{1}{\triangle_{i,j}}$ (monotonically decreasing with geographical distance). For simplicity, we assume that nodes have homogeneous traffic distributions. After normalization ($\sum_j T_{i,j}$=1), we have ($\ln n \gg \ln 2$):

$$T_{i,j} = \frac{1}{2\ln\frac{n}{2}} \frac{1}{\triangle_{i,j}} \approx \frac{1}{2\ln n} \frac{1}{\triangle_{i,j}} \tag{1}$$

Eq. 1 shows the desired traffic distribution model II. We also consider two different models (model I where $T_{i,j} \propto 1$, and model III where $T_{i,j} \propto e^{-\triangle_{i,j}}$) in Section 4, since they represent different node communication patterns. The movement in model I has no locality, meaning that people (e.g., businessmen) equally visit nearby and remote locations. On the other hand, the movement in model III is locality-oriented, meaning that people (e.g., farmers) only move around their home locations. Model II is in the middle. Compared to existing traffic distribution models in [21], [22], our models are more simple and abstracted. They describe the asymptotic traffic distribution with respect to the geographical distance.

Flows are only allowed to be routed among connected nodes (through links or chords). Flow interactions occur in nodes along a routing path from source to destination, which is a sequence of links and chords. The routing path is the shortest one determined by a *greedy-coin-change-based routing scheme*. In this routing scheme, each node greedily forwards the interaction traffic flow to a connected relay that is geographically closest to the destination [6], [23]. This flow routing scheme is used, since people usually make local decisions for their movements without a global view. The routing path (representing the movement trace) from $N_i$ to $N_j$ is denoted as $P_{i,j}$, with its length being $|P_{i,j}|$. We define the average number of routing hops as:

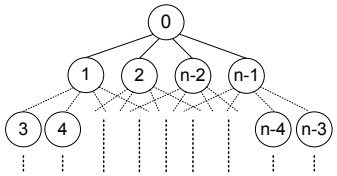


Fig. 5. An illustration of the 4-ary tree.

$$H = \frac{\sum_{i,j} T_{i,j}|P_{i,j}|}{\sum_{i,j} T_{i,j}} \tag{2}$$

$H$ represents the communication convenience (a smaller $H$ means more convenient communications). Our goal is to design a structure that can minimize $H$ without epidemic outbreaks. Then, we have the following observation:

***Proposition 1.*** If the network structure is isomorphic and all the nodes in the network have out-degrees of $D$, then the average traffic received by a node is $H$. The average amount of traffic going through each edge is $H/D$.

*Proof:* The interaction traffic from $N_i$ to $N_j$ is $T_{i,j}$, with a path length $|P_{i,j}|$. Each node on the path $P_{i,j}$, except for the source node $N_i$, receives this traffic. Therefore, $|P_{i,j}|$ nodes receive a traffic of $T_{i,j}$, and the total traffic is

$$\sum_{i,j} T_{i,j}|P_{i,j}| = \sum_{i,j} T_{i,j} \times \frac{\sum_{i,j} T_{i,j}|P_{i,j}|}{\sum_{i,j} T_{i,j}} = nH \tag{3}$$

Since the desired interaction traffic of each node is a specified unit, we have $\sum_{i,j} T_{i,j}$=$n$. Therefore, the average traffic received by a node is $H$. Note that the traffic of a node is shared by $D$ edges, and thus the average amount of traffic going through each edge is $H/D$. ■

Proposition 1 means that longer routing paths bring heavier forwarding traffic on each node. If the ring network has fewer chords, then people need to go through more locations to reach their destinations. Meanwhile, we have:

***Theorem 1.*** In any network structures, if all the nodes have out-degrees no more than $D$, then $H \in \Omega(\log_D n)$.

*Proof:* Since the out-degree of each node is bounded by $D$, a node can reach at most $D^{|P|}$ nodes through a path no longer than $|P|$. When the network structure is a $D$-ary tree, the root node can reach the most nodes with a limited path length. The lower bound of $H$ is that each node can be regarded as a root of a $D$-ary tree. Let us look into a specified node, $N_0$, as the root node. The neighbors of $N_0$ in the ring are put into the higher layers of the tree, and the remote nodes are assigned into the lower layers of the tree, as shown in Fig. 5. This is because $N_0$ has more interaction traffic to nearby nodes than remote nodes. In Fig. 5, nodes $N_1, N_2, N_{n-1}, N_{n-2}$ are placed into the second layer, while the interaction traffic between the root and these nodes are $T_{0,1} = T_{0,n-1} = \frac{1}{2\ln n}\frac{1}{1}$ and $T_{0,2} = T_{0,n-2} = \frac{1}{2\ln n}\frac{1}{2}$. Then, we have the following inequality:

$$\begin{aligned} H &= \frac{\sum_i T_{0,i}|P_{0,i}|}{\sum_i T_{0,i}} = \sum_{i=1}^{n-1} T_{0,i}|P_{0,i}| = 2 \times \sum_{i=1}^{\frac{n}{2}-1} T_{0,i}|P_{0,i}| \\ &\geq \frac{1}{\ln n} \sum_{x=1}^{\log_D n} \left[ x \times \int_{(D^{x-1}+\ldots+D^0)/2}^{(D^x+\ldots+D^0)/2} \frac{1}{y} \mathrm{d}y \right] \end{aligned} \tag{4}$$

This is because the tree has $O(\log_D n)$ layers. For the $x^{th}$ layer, the tree contains $D^{x-1}$ nodes. Then, we have

$$\int_{(D^{x-1}+...+D^0)/2}^{(D^x+...+D^0)/2} \frac{1}{y}\mathrm{d}y = \ln \frac{D^x+...+D^0}{D^{x-1}+...+D^0} > \ln D \quad (5)$$

Therefore, Eq. 4 can be rewritten as

$$H > \frac{1}{\ln n}\left[\sum_{x=1}^{\log_D n} x \ln D\right] \approx \frac{\log_D^2 n \ln D}{2 \ln n} = \frac{1}{2}\log_D n \quad (6)$$

where $\log_D n = \frac{\ln n}{\ln D}$. Hence, we have $H \in \Omega(\log_D n)$. ■

### 2.5 The Infection Rate Model

We start with the infection rate model B, which is a simplification of the classic SIS model [4]. The eigenvalue approach [24] is not used because of complexity. In our model, the infection rate can be traffic-sensitive. We highlight that a desired traffic flow with *different routing paths* have different impacts on epidemic spreadings [24]. This paper uniquely connect the traffic flow routing and the epidemic spreading through the network structure design.

Flows have states of being susceptible or infected. A susceptible flow represents a group of moving people who do not have the disease, but can be infected. A infected flow represents a group of moving people, who have the disease and can spread the disease at the locations they go through. Infected flows can go back into the susceptible state upon recovery, and can be reinfected. For isomorphic structures, Theorem 1 shows that the average traffic per edge is $H/D$. In the infection rate model B, let us consider a node with $D$ incoming edges, each of which loads a traffic of $H/D$. If the flows on one incoming edge are infected, then they would bring an infection rate (the infection probability per time unit) of $\lambda\sqrt{H/D}$, where $\lambda$ is a coefficient ($\lambda > 0$). The square root depicts the decreasing hazard infection of subsequent interactions [25]. Other sub-linear functions can also be used here. Section 4 further considers a constant rate model of $\lambda$ as model A, and a linear rate model of $\lambda H/D$ as model C. While model A is traffic-insensitive, model C is traffic-sensitive. Model B is in the middle.

Based on the existing literature [4], $\lambda$ and $r$ are fixed. Let $f(t)$ and $g(t)$ (or simply $f$ and $g$) denote the average fractions of edges that load infected and susceptible flows at time $t$, respectively. Each node has, on average, $Df$ incoming edges that load infected flows. Hence, the infection probability for susceptible outgoing flows per time unit is:

$$1-(1-\lambda\sqrt{\frac{H}{D}})^{Df} \approx \lambda\sqrt{HD}f = \mu f \quad (7)$$

In Eq. 7, we define $\mu = \lambda\sqrt{HD}$ for presentation simplicity. We assume $\lambda\sqrt{H/D} \ll 1$. Otherwise, the epidemic outbreak cannot be controlled. Hence, a total fraction, $\mu fg$, of all flows are infected during a time unit. Let $r$ denote the recovery rate of infected flows. Therefore, we have:

$$\frac{\mathrm{d}g}{\mathrm{d}t} = rf - \mu fg \quad \text{and} \quad \frac{\mathrm{d}f}{\mathrm{d}t} = \mu fg - rf \quad (8)$$

Based on the existing literature [4], the solution to Eq. 8 is:

$$f(t) = f(0)\frac{(\mu-r)e^{(\mu-r)t}}{\mu - r + \mu f(0)e^{(\mu-r)t}} \quad (9)$$

TABLE 2
Notations in this paper

| Notation | Description |
|---|---|
| $N_i$ | Node with an ID of $i$ in the ring. |
| $n$ | Total number of nodes in the ring. |
| $D$ | In-/Out-degrees of all nodes (only used for isomorphic structures). |
| $d$ | Number of chords for half a ring. We have $d = D/2-1$. |
| $\triangle_{i,j}$ | Geographical distance between $N_i$ and $N_j$. |
| $T_{i,j}$ | Desired interaction traffic from $N_i$ to $N_j$. |
| $P_{i,j}$ | Routing path from $N_i$ to $N_j$. $\lvert P_{i,j}\rvert$ denotes the path length. |
| $T_l$ | Total interaction traffic going through $N_l$. |
| $H$ | Average number of routing hops. |
| $f$ $(g)$ | Fraction of edges that load infected (susceptible) flows. |
| $\lambda, \mu$ | Coefficients for infection rate models. |
| $r$ | Recovery rate of the infected flows. |
| $Q(D)$ | Fraction of nodes with in-degree $D$ (only used for polymorphic structures). |
| $\langle\cdot\rangle$ | Mean value of the corresponding variable. |

If $\mu < r$, $f(t)$ decreases exponentially, meaning that the epidemic outbreak is controlled. Hence, we have:

$$\mu < r \quad \text{or} \quad \sqrt{HD} < \frac{r}{\lambda} \quad (10)$$

A larger degree $D$ generally brings a smaller $H$, since more chords can shrink the network diameter. Eq. 10 means that short routing paths can contribute to the control of epidemic outbreaks. This is because flows visit fewer locations, when $H$ is smaller. However, the pattern of Eq. 10 is very complex:

- If $H$ logarithmically decreases with $D$, the constraint of controlling epidemic outbreaks is equivalent to a degree limitation. For example, if $H = \ln n/\ln D$, then $\sqrt{HD}$ increases monotonically with $D > 2.718$, i.e., $D$ should be smaller than a threshold. Our goal becomes minimizing $H$ with a limited degree.
- On the other hand, if $H$ exponentially decreases with $D$, then a larger $D$ is better, which is counter-intuitive. For example, if $H = n/2^D$, then $\sqrt{HD}$ decreases monotonically with $D > 1.443$.

The above phenomena shows the tradeoff between local and global infections. If $D$ is small, then the traffic will aggregate along local nodes ($H$ is large). In this case, susceptible flows going through locations with infected flows are very likely to be infected, resulting in a local infection phenomenon. On the other hand, if $D$ is large, then the traffic is spread widely through chords. At this time, an infected flow can reach many susceptible remote flows to spread the disease, leading to the global infection phenomenon. The challenge is that $H$ cannot be directly derived as a function of $D$. This is because networks with the same $D$ can have a different $H$ due to different network structures.

### 2.6 Problem Formulation

This paper studies a network structure design problem constrained by the epidemic outbreaks. The network structure refers to the number and distribution of chords on the ring. Our objective is to minimize the average number of routing hops (minimize $H$), and the constraint is that epidemics will not outbreak. The variable of our problem is the number

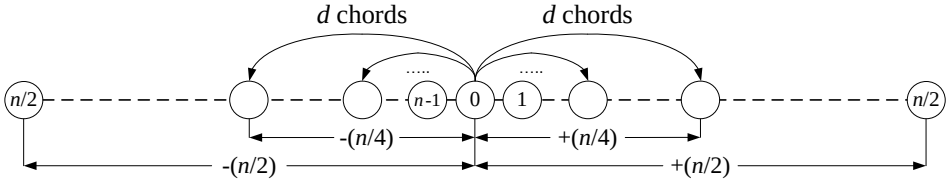



Fig. 6. An illustration of the binary-cut network structure (only the chords of $N_0$ are shown).

and distribution of chords (how nodes connect each other on the ring; as a result, the degree $D$ is a variable). The distribution of the desired traffic ($T_{i,j}$) is given. The routing scheme (given the number and distribution of chords, which path does $T_{i,j}$ route from $N_i$ to $N_j$) is given. The epidemic outbreak constraint ($\sqrt{HD} < \frac{r}{\lambda}$ for infection rate model B) is given. In general, more chords brings a smaller $H$, but the network may be more vulnerable to epidemic outbreaks ($\sqrt{HD}$ may be larger, since the degree $D$ is larger). Networks with the same number of chords can have a different $H$ due to different chord distributions.

Our problem stands for an effective network quarantine with minimal restrictions on communication activities. The ring network is an abstraction of the real-world geographic map, as shown in Fig. 1. The quarantine rule is represented by the number and distribution of chords. Two unconnected nodes (no chord) mean that they are isolated from each other in the quarantine, in order to control epidemic outbreaks. A real-world example is that Australia clamped down on the entry of individuals traveling from West Africa [3]. Australia and West Africa are represented as two nodes on the ring, while there is no chord connecting these two nodes. Such a quarantine rule is executed by the Australian Customs and Border Protection Service. The average number of routing hops represents the convenience level of the communication activities. A smaller $H$ means that fewer restrictions are put on the communication activities. The epidemics in this paper include human diseases, as well as epidemic information propagations in distributed systems (more detailed descriptions can be found in [2]).

In the next section, we design isomorphic structures, under the desired traffic distribution model II and the infection rate model B. In Section 4, we discuss other traffic distribution models and infection rate models. The reason for the isomorphic structures is that networks with heterogenous degrees are more vulnerable to epidemics [26]. We also show a greedy solution for constructing polymorphic structures in Section 5. Finally, all the notations are shown in Table 2.

## 3 Isomorphic Structure Design

In this section, we design two hypercube-based isomorphic structures that correspond to property (6) in Table 1. Since the structure is isomorphic, we only describe the chord distribution of node $N_0$ for clear presentation. A chord from the node $N_0$ to $N_i$ ($i<\frac{n}{2}$) indicates that $N_0$ has a jump size of $i$. The out-degree of $N_0$ is $D$, including two links and $D-2$ chords. We use symmetric designs for each side of the ring, meaning that $d = \frac{D}{2} - 1$ chords are assigned for the left-side (or right-side) half ring of $N_0$. The core idea of our structure design is to find a good tradeoff between the degree and the average number of routing hops, while the greedy-coin-change-based routing is well-supported. All of these factors can be found in hypercube structures [5], [7] because of their structural symmetries. Epidemic outbreaks are revealed by the tradeoff between the local and global infections, which are, in turn, controlled by the degree.

First, we discuss the hypercube-based *binary-cut* structure, which is shown in Fig. 6. In this network structure, the jump sizes are $\{+\frac{n/2}{2}, +\frac{n/2}{2^2}, ..., +\frac{n/2}{2^d}, +1\}$ for the right-side half ring and $\{-\frac{n/2}{2}, -\frac{n/2}{2^2}, ..., -\frac{n/2}{2^d}, -1\}$ for the left-side half ring. For example, if $n = 16$ and $d = 2$ ($D = 6$), $N_0$ would have chords to nodes $\{N_2, N_4, N_{12}, N_{14}\}$ and links to nodes $\{N_1, N_{15}\}$. We have assumed $d < \log_2 n$ such that nodes in the range $(-\frac{n/2}{2^d}, +\frac{n/2}{2^d})$ are archived through step-by-step links (jump sizes of 1). Then, the bounds of $H$ for the binary-cut structure are:

***Theorem 2.*** In the binary-cut structure, the bounds of $H$ are $H \in O(\frac{3^d n}{4^d \ln n})$ and $H \in \Omega(\frac{n}{2^d \ln n})$.

The proof of Theorem 2 is attached in the supplemental material. Asymptotically, $HD$ monotonically decreases with $D$, meaning that a larger degree is better. This structure can support at most $d = \log_2 \frac{n}{2} - 1$ chords for half a ring. The jump sizes for half a ring are $\{+1, +2, +4..., +\frac{n/2}{2}\}$.

However, the binary-cut structure fails to consider the desired traffic distribution for optimizing $H$. While each node has more traffic to its geographically-nearby nodes, we should provide more jumps to geographically-nearby nodes and fewer jumps to geographically-remote nodes. Following this intuition, we can improve the binary-cut network structure to the *binary-traffic-cut* network structure. Instead of jumping to the middle node of each interval, we jump to the node that forwards *half of its desired traffic*. Considering that we have $\int_1^{\sqrt{n}} \frac{1}{i} \mathrm{d}i \approx \int_{\sqrt{n}}^{n} \frac{1}{i} \mathrm{d}i$, the jump sizes are set to be $\{+\sqrt{n/2}, +\sqrt[4]{n/2}, +\sqrt[8]{n/2}, ..., +\sqrt[2^d]{n/2}, +1\}$ and $\{-\sqrt{n/2}, -\sqrt[4]{n/2}, -\sqrt[8]{n/2}, ..., -\sqrt[2^d]{n/2}, -1\}$. Here, the desired traffic from node $N_0$ to the nearest $\sqrt{n/2}$ nodes is the same as its traffic to the remaining nodes. The bounds of $H$ for the binary-traffic-cut structure are:

***Theorem 3.*** In the binary-traffic-cut structure, the upper and lower bounds of $H$ are $H \in O(\frac{\sqrt{n}}{\ln n} + \frac{\ln^d n}{2^{d^2}} n^{\frac{1}{2^d}})$ and $H \in \Omega(\frac{\sqrt{n}}{\ln n} + \frac{1}{\ln n} n^{\frac{1}{2^d}})$, respectively.

The proof of Theorem 3 is attached in the supplemental material. The binary-traffic-cut structure can support at most $O(\ln \ln n)$ chords. However, even if $d$ is very small, the first part of $\frac{\sqrt{n}}{\ln n}$ dominates the bounds of $H$. For this structure, $HD$ monotonically increases with $D$, meaning that we need to control the degree to avoid epidemic outbreaks:

***Corollary 1.*** For a binary-traffic-cut network structure without epidemic outbreaks, we have $H \in O(\frac{\sqrt{n}}{\ln n} + \frac{\ln^d n}{2^{d^2}} n^{\frac{1}{2^d}})$ and $H \in \Omega(\frac{\sqrt{n}}{\ln n} + \frac{1}{\ln n} n^{\frac{1}{2^d}})$, where $d \approx (r^2 \ln n)/(2\lambda^2 \sqrt{n})$.

If $(r^2 \ln n)/(2\lambda^2\sqrt{n})<1$, epidemic outbreak is inevitable for this structure. In terms of minimizing $H$, the binary-traffic-cut structure outperforms the binary-cut structure, since it considers the desired traffic distribution. However, in the view of minimizing $H$ without epidemic outbreaks, it is hard to analytically compare these two structures, since they have different vulnerabilities to epidemic outbreaks.

## 4 Other Traffic Distribution and Infection Rate Models for Isomorphic Structures

In this section, we study isomorphic structures with different traffic distribution and infection rate models.

### 4.1 Other Traffic Distribution Models

In previous sections, we only study the desired traffic distribution model II, where $T_{i,j} \propto 1/\triangle_{i,j}$. In this subsection, two different traffic distribution models (model I and model III) are discussed. These three models present different node communication patterns. Compared to model II (i.e., the most general case), models I and III are two more extreme models that have simpler properties as follows.

Let us start with the desired traffic distribution model I, where we have a uniform distribution of $T_{i,j} = \frac{1}{n}$. In this case, the binary-cut structure is equivalent to the binary-traffic-cut structure. For each jump size on one side of the ring, half of the flows to nodes on that side will take it. For example, if $n = 16$ and $d = 2$, flows from $N_0$ to $\{N_4, N_5, N_6, N_7\}$ include a jump size of $+4$, flows from $N_0$ to $\{N_2, N_3, N_6, N_7\}$ include a jump size of $+2$, and flows from $N_0$ to $\{N_1, N_3, N_5, N_7\}$ include a jump size of $+1$. Therefore, the average path length through chords is $\Theta(d)$. For each node in the interval of $(0, \frac{n}{2^d})$, it is achieved through links (jump sizes of 1), the average path length of which is $\Theta(\frac{n}{2^d})$. Therefore, we have $H \in \Theta(d + \frac{n}{2^d})$, where $d \in \Theta(\log_2 n)$ is the best choice. If $HD < r^2/\lambda^2$, we have $H \in \Theta(\log_2 n)$ for the binary-cut structure. However, better structures may exist. For the same number of chords, the Ulysses butterfly network structure [9] achieves a lower network diameter than does the hypercube structure.

The desired traffic distribution model III has an exponential traffic distribution of $T_{i,j} \propto e^{-\triangle_{i,j}}$. Note that, the distribution in model III decays exponentially, and thus the traffic to remote nodes can be ignored (people's movements are extremely locality-oriented). Assuming $N_0$ is the source, then each node can be achieved in, at most, $\frac{n}{2}$ steps (by links), and at least 1 step (by a chord). Then, we have the following inequation:

$$\frac{n}{2}\times \int_{2\ln\frac{n}{2}}^{\frac{n}{2}} e^{-i}\mathrm{d}i < \frac{n}{2}\times e^{-2\ln\frac{n}{2}} \ll 1\times \int_{1}^{2\ln\frac{n}{2}} e^{-i}\mathrm{d}i \quad (11)$$

Eq. 11 means that the nearest $2\ln\frac{n}{2}$ nodes on one side of the ring are more important, while the traffic from $N_0$ to the remaining nodes on that side can be ignored. Even if all the other nodes can only be achieved by links, their influences on $H$ are much smaller than the nearest $2\ln\frac{n}{2}$ nodes. Therefore, all the chords are only necessary to connect the nearest $O(\ln n)$ nodes, regardless of the remaining remote nodes. The insight is that we need smaller jump sizes for a faster traffic decay, where remote connections are useless. This result is stated as the property 4 in Table 1.

### 4.2 Other Infection Rate Models

In previous sections, we use the infection rate model B, where $\mu = \lambda\sqrt{HD}$. The reason behind the sub-linearity of model B is that people are more likely to be infected by the initial interactions with infectors than the subsequent interactions with infectors. In this subsection, two different infection rate models (model A and model C) are discussed. These three infection rate models present different epidemic infectivities. Compared to model B (i.e., the most general case), models A and C are two more extreme models that have simpler properties as follows.

First, we consider the infection rate model A, where the infection rate is a constant, i.e., traffic-insensitive. In this case, Eq. 7 can be rewritten as (let $\mu = \lambda D$ for consistency):

$$1 - (1-\lambda)^{Df} \approx \lambda D f = \mu f \quad (12)$$

Eqs. 8 and 9 remain the same, while Eq. 10 changes to:

$$D < \frac{r}{\lambda} \quad (13)$$

Eq. 13 means that the constraint of controlled outbreaks is equal to the degree limitation. The insight is that the global infection dominates, and thus we need to restrict the connections among different locations. This result is stated as the property 1 in Table 1. Under the traffic distribution model I and the infection rate model A, our problem is reduced to minimizing $H$ with a limited degree. As shown in [9], Ulysses butterfly network structure is asymptotically optimal for this problem, where we have $H \in \Omega(\log_{r/\lambda} n)$. This result is stated as property 5 in Table 1.

As for the infection rate model C, we consider that the infection rate is traffic-sensitive. For a given location, each incoming edge that loads infected flows would independently bring an infection probability that is linearly proportional to the traffic on that edge. Then, Eq. 7 can be rewritten as (let $\mu = \lambda H$ for presentation consistency):

$$1 - (1 - \lambda\frac{H}{D})^{Df} \approx \lambda H f = \mu f \quad (14)$$

Eqs. 8 and 9 remain the same, while Eq. 10 changes to:

$$H < \frac{r}{\lambda} \quad (15)$$

Eq. 15 means that the network structure with a smaller $H$ can resist epidemics better, and thus the fully-connected network is the best choice. This is because the local infection becomes the major factor for epidemic outbreaks. It implies that the communication restrictions lead to negative effects, under the infection rate model C. In this case, one node should connect to as many nodes as possible, in order to mitigate the traffic aggregation on local links. This result is stated as property 3 in Table 1.

## 5 Polymorphic Structure Design

In this section, we study polymorphic structures, under the desired traffic model I and the infection rate model A. Other models are not explored due to their complexities.

**Algorithm 1** Greedy Construction

**Input:** A virtual ring network that has no chord;
**Output:** A polymorphic network structure;

1: **while** $\frac{\langle D^2\rangle}{\langle D\rangle} < \frac{r}{\lambda}$ **do**
2: **for** each pair of unconnected nodes **do**
3: Calculate the corresponding $\Delta H$ and $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$, if a chord is added for this pair of nodes;
4: Add the chord with the highest $\Delta H/\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$;
5: Remove the last added chord to guarantee $\frac{\langle D^2\rangle}{\langle D\rangle} < \frac{r}{\lambda}$;
6: **return** the current network structure;

**Algorithm 2** Greedy Dismantlement

**Input:** A fully-connected virtual ring network;
**Output:** A polymorphic network structure;

1: **while** $\frac{\langle D^2\rangle}{\langle D\rangle} > \frac{r}{\lambda}$ **do**
2: **for** each pair of nodes connected by a chord **do**
3: Calculate the corresponding $\Delta H$ and $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$, if a chord is removed for this pair of nodes;
4: Remove the chord with the lowest $\Delta H/\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$;
5: **return** the current network structure;

### 5.1 Contributions of an Additional Chord

This subsection studies how $H$ reduces, if an additional chord is added into the current ring network (under the desired traffic model I). Let $\Delta H$ denote the amount of reduced $H$. Let us start with a virtual ring network that has no chord, and then consider adding a chord from $N_{i-x}$ to $N_i$ with a jump size of $x$. Then, nodes $\{N_{i-n/2}, ..., N_j, ..., N_{i-x}\}$ can benefit from this chord. For $N_j$ specifically, its routing paths of flows to nodes $\{N_i, ..., N_{j+n/2}\}$ are shortened by $x-1$. Under the desired traffic model I, we have:

$$\Delta H = (x-1)\sum_{j=0}^{\frac{n}{2}-x}(\frac{n}{2}-x-j) \approx \frac{1}{2}(\frac{n}{2}-x)^2 x \qquad (16)$$

Eq. 16 is maximized, when we use a chord with the jump size of $x = \frac{n}{6}$. The insight behind Eq. 16 is the tradeoff between (1) the number of routing paths that can benefit from the chord and (2) the saved lengths of routing paths brought by the chord. Here, we do not further explore the contribution of an additional chord on a ring with multiple chords. However, note that the $\Delta H$ brought by an additional chord can be calculated in polynomial time. This is because we have $O(n^2)$ pairs of nodes, while the routing path for each pair of nodes can be determined within $O(n)$.

### 5.2 Epidemics in Polymorphic Structures

This subsection introduces an advanced epidemic model [26] for polymorphic structures under the infection rate model A. To capture the structural heterogeneity, let $Q(D)$ denote the fraction of nodes with in-degree $D$, and let $f_D(t)$ denote the fraction of incoming edges that load infected flows in locations with in-degree $D$ at time $t$. For the model A of the constant infection probability, we have:

$$\frac{\mathrm{d}f_D(t)}{\mathrm{d}t} = \lambda D[1-f_D(t)]\Theta(f(t)) - r f_D(t) \qquad (17)$$

Eq. 17 is similar to Eq. 8, which is for isomorphic structures. The fraction of outgoing edges that load susceptible flows in locations with in-degree $D$ is $[1-f_D(t)]$. $\Theta(f(t))$ is the total fraction of edges that load infected flows. Therefore, the first term, $\lambda D[1-f_D(t)]\Theta(f(t))$, indicates the fraction of edges that load new infected flows at locations with in-degree $D$. The last term, $r f_D(t)$, shows the recovery. Existing work in [26] shows the following result:

$$\frac{\langle D^2\rangle - \langle D\rangle^2}{\langle D\rangle} + \langle D\rangle < \frac{r}{\lambda} \qquad (18)$$

$\langle D\rangle$ is average node degree, and $\langle D^2\rangle - \langle D\rangle^2$ represents the node degree variance. Eq. 18 depicts the prerequisite of controlled outbreaks in polymorphic structures, under the infection rate model A. In contrast, the constraint for isomorphic structures are shown in Eq. 13. The insight of Eq. 18 is that a larger degree variance also brings a more vulnerable network. Under the infection rate model A, both the average degree and the degree variance determines the network resistance to epidemics. $\frac{\langle D^2\rangle}{\langle D\rangle}$ indicates the network vulnerability to epidemics (the larger, the more vulnerable).

### 5.3 Constructing Polymorphic Structures

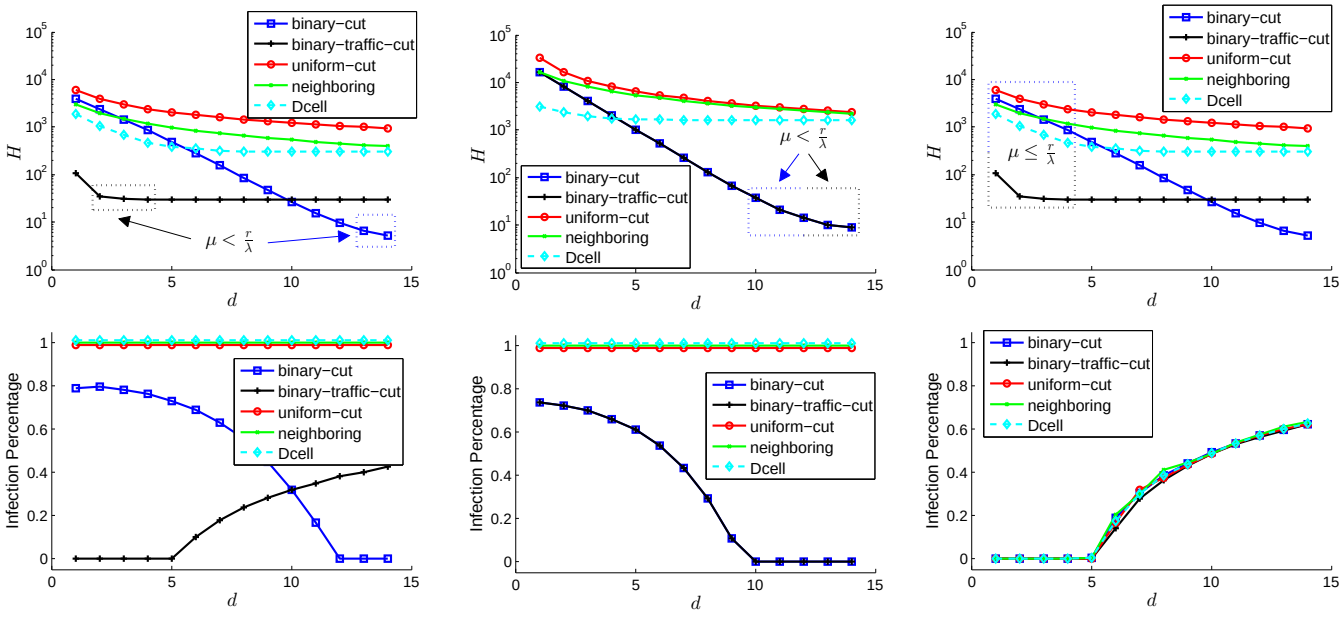
This subsection describes two greedy algorithms for polymorphic structures. Algorithm 1 starts with a ring network that has no chords, and then iteratively adds chords with the consideration of (1) $\Delta H$ that indicates the reduced average number of routing hops (the benefit of that chord) and (2) $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$ that represents the increased network vulnerability (the cost of that chord). Algorithm 1 iteratively adds the chord with the highest ratio of $\Delta H$ to $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$ (the benefit-to-cost ratio). In the event of a tie, a random one is picked. A chord with a large $\Delta H$ means that it can effectively minimize the number of average hops. A chord with a small $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$ indicates that adding this chord only slightly increases the network vulnerability to epidemics. In contrast, Algorithm 2 starts with a fully-connected ring network, and then iteratively removes the chord with the lowest ratio of $\Delta H$ to $\Delta\big[\frac{\langle D^2\rangle}{\langle D\rangle}\big]$. Algorithms 1 and 2 are extendable to heterogeneous traffic distribution and epidemic spreading models [27], [28], since they construct polymorphic network structures based on the benefit-to-cost ratio of each chord. However, both Algorithms 1 and 2 are suboptimal. They may be trapped into local optima due to the greedy nature. The optimal solution remains to be explored.

## 6 Experiments

This section conducts experiments to evaluate the proposed structures. Codes are published in [29].

### 6.1 Experiments on Ring Networks

**Settings.** In this subsection, we focus on the proposed isomorphic structures in a ring network. The ring network has $n = 2^{17}$ nodes, with the number of chords ranging from 2 to 30. The infection coefficient $\lambda$ is 0.01, while the recovery rate $r$ varies in different settings, in order to observe the epidemic outbreak point. $1\%$ of the total edges




(a) $T_{i,j} \propto \frac{1}{\triangle_{i,j}}$, $\mu = \lambda\sqrt{HD}$, $r = 0.15$ (b) $T_{i,j} = \frac{1}{n}$, $\mu = \lambda\sqrt{HD}$, $r = 0.3$ (c) $T_{i,j} \propto \frac{1}{\triangle_{i,j}}$, $\mu = \lambda D$, $r = 0.1$

Fig. 7. The evaluation results. The top line figures show the relationship between $d$ and $H$, and the bottom line figures show the relationship between $d$ and the infection percentage of nodes after $1,000$ time units. (a) The evaluation results under the desired traffic model II and the infection rate model B, with a recovery rate $r = 0.15$. (b) The evaluation results under the desired traffic model I and the infection rate model B, with a recovery rate $r = 0.3$. (c) The evaluation results under the desired traffic model II and the infection rate model A, with a recovery rate $r = 0.1$.

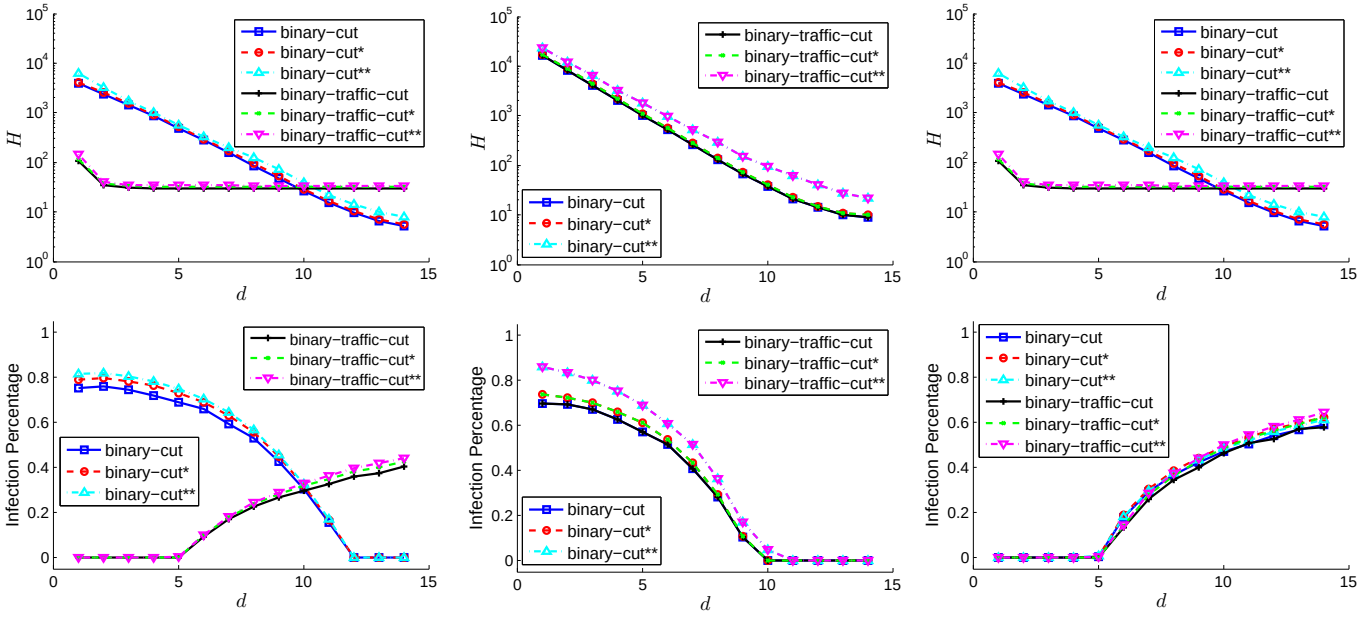
are initialized as being infected, while the percentage of infectors is checked again after $1,000$ time units. The time period of $1,000$ time units is long enough to guarantee that the infection percentage becomes steady. We do not consider the other initializations on the percentage of infectors, since [4] has shown that the initialized percentage of infectors has a very limited influence on the eventual epidemic outbreak (or not). Here, the desired traffic model III is not considered, since all chords tend to connect to the nearest nodes. The infection rate model C is not considered, since it results in a fully-connected network as an optimal solution. Therefore, we focus on the desired traffic model I ($T_{i,j} = \frac{1}{n}$) and model II ($T_{i,j} \propto \frac{1}{\triangle_{i,j}}$). As for the infection rate models, we only pay attention to model A ($\mu = \lambda D$) and model B ($\mu = \lambda\sqrt{HD}$).

Three additional isomorphic structures are used for comparison. The first one is the uniform-cut structure, where the chords uniformly cut the half-ring. In this structure, the jump sizes are $\{+1, +\frac{n}{2d}, +\frac{2n}{2d}, ..., +\frac{(d-1)n}{2d}\}$ for half a ring. The second one is the neighboring structure, where each node connects to the nearest $D$ geographical neighbors. The third one is the Dcell structure [30], which was proposed for the connections among servers in data center networks.

**Testing Different Structures.** The experimental results are shown in Fig. 7. The top-line figures show the relationship between $d$ and $H$, while the bottom-line figures show the relationship between $d$ and the percentage of infected flows after $1,000$ time units. The three subfigures in Fig. 7 describe three different settings of traffic distribution models and infection rate models. The top lines of Figs. 7(a) and 7(c) are the same, since they have the same traffic model. In Fig. 7(b), the performance of the binary-cut and binary-traffic-cut structures are the same, due to the uniform traffic model. In the top line figures, the dashed boxes show the theoretical interval of controlled epidemic outbreaks (i.e., the theoretical constraint of $\mu < \frac{r}{\lambda}$). The real infection percentages, after $1,000$ time units (i.e., epidemic outbreaks or not), are shown in the bottom line of Fig. 7.

It can be seen that the uniform-cut and the neighboring structures are useless, since they cannot control epidemic outbreaks. They have very large average numbers of routing hops. Although Dcell has a decent average number of routing hops, it cannot control epidemic outbreaks. For the desired traffic model II and the infection rate model B in Fig. 7(a), the binary-cut structure controls outbreaks with a large $d$, and the binary-traffic-cut structure controls outbreaks with a small $d$. This is consistent with our theoretical results. The binary-cut structure can achieve the smallest $H$ without outbreaks. For the desired traffic model I and the infection rate model B in Fig. 7(b), the performance of the two proposed structures are the same, and they can control epidemic outbreaks when $d$ is large enough. The desired traffic model II and the infection rate model A in Fig. 7(c) show that the epidemic outbreaks only depend on the degree. These results verify that the prerequisite of controlled epidemic outbreaks (i.e., $\mu < \frac{r}{\lambda}$) is accurate.

**Sensitivity Experiments.** One step further, here we also study the sensitivities of the proposed structures, since the assumption of isomorphic structures may be too strong for real-world applications. We want to check whether this assumption can be relaxed or not. Hence, we randomly rewire a fraction of chords for the proposed structures to check their sensitivities. If a chord is rewired, it will reconnect to a pair of nodes that are randomly selected. For the binary-cut and binary-traffic-cut structures, the resulting structures with 10% rewired chords are denoted as binary-cut* and binary-traffic-cut*, respectively. Then, the resulting




(a) $T_{i,j} \propto \frac{1}{\triangle_{i,j}}, \mu = \lambda\sqrt{HD}, r = 0.15$ (b) $T_{i,j} = \frac{1}{n}, \mu = \lambda\sqrt{HD}, r = 0.3$ (c) $T_{i,j} \propto \frac{1}{\triangle_{i,j}}, \mu = \lambda D, r = 0.1$

Fig. 8. The sensitivity experiments. The settings are the same as those in Fig. 7. The binary-cut* and binary-traffic-cut* denote binary-cut and binary-traffic-cut structures with 10% randomly rewired chords, respectively. Similarly, The binary-cut** and binary-traffic-cut** denote binary-cut and binary-traffic-cut structures with 50% randomly rewired chords, respectively.

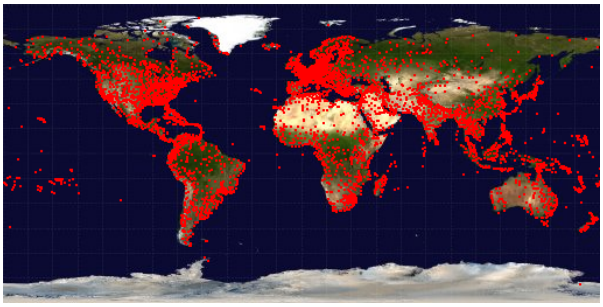


Fig. 9. The geographic map of the real data-driven experiment.

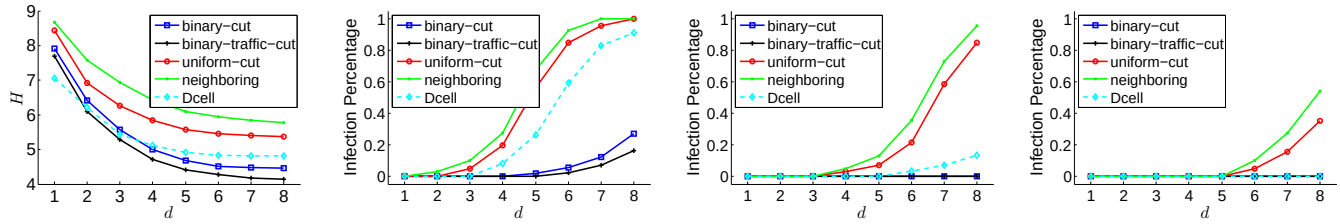
structures with 50% rewired chords are denoted as binary-cut** and binary-traffic-cut**, respectively. The sensitivity experiment results are shown in Fig. 8. It can be seen that, rewired chords have a very limited impact, in terms of both the average number of routing hops ($H$) and the infection percentage. The network traffic model and infection rate model also do not significantly change the result. The average number of routing hops ($H$) slightly increases, since the hypercube-based structures are greatly fault-tolerant with respect to greedy-coin-change-based routings. The variance of the infection percentage is also limited, since the isomorphic structure is not destructed too much. Overall, these two structures are not sensitive to a small portion of randomly rewired chords, i.e., the isomorphism assumption can be relaxed for real-world applications.

### 6.2 Real Data-Driven Experiments

**Settings.** This subsection conducts real data-driven experiments to verify the applicability of our approach. We use a real airline dataset of OpenFlights [31], which contains 59,036 routes between 3,209 airports on 531 airlines spanning the globe, as shown in Fig. 9 (airports are marked as red points). Airports correspond to nodes in our model. We discretize the Earth through 100 meridians and 100 parallels (10,000 points). Each airport is rounded to its nearest geographic point. The existing airline routes imply the information on the desired traffic distribution. In other words, the desired traffic distribution model is given, where each flight stands for a unit traffic. If we use the existing airline routes as chords, then the average number of routing hops ($H$) is one, but this structure fails to control epidemic outbreaks. To apply our ring model, X-Y routing scheme is used for node communications. The routing has two stages: the first stage routes along the meridian ring of the source node, and the second stage routes along the parallel ring of the destination node. An example of such a routing has been shown in Fig. 4. The number of routing hops is the sum of these two stages on two different rings.

We use the real Ebola dataset in three countries (Guinea, Liberia, and Sierra Leone) reported by Word Health Organization [32] to model epidemic spreadings. Table 3 shows the data statistics, in terms of the number of recovers and deaths. The number of recovers are significantly smaller than the number of deaths in Guinea, approximately equal to the number of deaths in Liberia, and significantly larger than the number of deaths in Sierra Leone. The fractions of recovers and deaths are used as the recover and infection rates, respectively. Here, the infection rate is traffic-insensitive (infection rate model A). The other settings are the same as those in the previous subsection.

**Experimental results.** The results for the real data-driven experiments are shown in Fig. 10. Fig. 10(a) shows the relationship between $H$ and $d$. A larger node degree leads to a smaller average number of routing hops. The binary-traffic-cut outperforms other structures, when $d$ is large. It




(a) Number of routing hops. (b) Epidemics in Guinea. (c) Epidemics in Liberia. (d) Epidemics in Sierra Leone.

Fig. 10. Evaluation results for real data-driven experiments.

TABLE 3
Ebola statistics reported by Word Health Organization.

| Dataset | Recover | Death | Total |
|---|---|---|---|
| Guinea (12/28/2015) | 1,268 | 2,536 | 3,804 |
| Liberia (5/9/2015) | 5,860 | 4,806 | 10,666 |
| Sierra Leone (11/7/2015) | 10,167 | 3,955 | 14,122 |

has an $H$ of about 4, when $d = 8$. In contrast, neighboring has the largest $H$ among all the structures. Another notable point is that Dcell has the smallest $H$ when $d$ is small (when $d = 1$). Fig. 10(b), Fig. 10(c), and Fig. 10(d) shows the eventual infection percentages under the dataset of Guinea, Liberia, and Sierra Leone, respectively. A larger $d$ always leads to a larger infection percentage. Note that, in our structure design, each airport can have four routes to nearby airports (on meridian and parallel), and $2d$ routes to remote airports ($d$ for meridian-remote airport and $d$ for parallel-remote airport). This is because the infection rate is not traffic-sensitive, and thus, we should cap the node degree to avoid epidemic outbreaks. Fig. 10(b) has the highest infection percentages, since the dataset of Guinea has the highest death-to-recover ratio, and the infection rate is larger than the recovery rate. In Fig. 10(b), uniform-cut and neighboring can only control epidemic outbreaks when $d = 2$ with $H \approx 7$. Dcell controls epidemic outbreaks when $d = 3$ with $H \approx 6$. Binary-cut and binary-traffic-cut control epidemic outbreaks when $d = 5$ with $H \approx 5$. In Fig. 10(c), uniform-cut and neighboring control epidemic outbreaks when $d = 4$ with $H \approx 6$. Dcell controls epidemic outbreaks when $d = 6$ with $H \approx 7$. Binary-cut and binary-traffic-cut control epidemics with $H \approx 4$. Fig. 10(d) shows similar results with Fig. 10(c), since Sierra Leone has the lowest death-to-recover ratio. Real data-driven experiments confirm the real-world applicability of our approach.

## 7 Conclusion

This paper studies network structures with minimum traffic flow routings, while controling epidemic outbreaks through regulating the number and distribution of chords on a ring network. The objective is to design a structure (quarantine rules) that can minimize the average number of routing hops, while the epidemic outbreaks are controlled. For isomorphic structures, we provide a systematic structure design on nine different cases. Two hypercube-based structures are explored. Sensitivity experiments demonstrate that the isomorphism assumption can be relaxed. For polymorphic structures, we provide a greedy solution. Our work casts new light on the effective network quarantine that places minimal restrictions on communication activities.

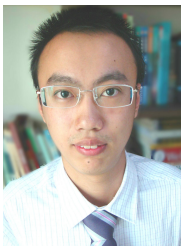
## Acknowledgments

This work is supported in part by NSF grants CNS 149860, CNS 1461932, CNS 1460971, CNS 1439672, CNS 1301774, ECCS 1231461, ECCS 1128209, and CNS 1138963.

## References

[1] C. Lagorio, M. Dickison, F. Vazquez, L. A. Braunstein, P. A. Macri, M. V. Migueles, S. Havlin, and H. E. Stanley, "Quarantine-generated phase transition in epidemic spreading," *Physical Review E*, vol. 83, no. 2, p. 026102, 2011.
[2] P. T. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massoulié, "Epidemic information dissemination in distributed systems," *Computer*, vol. 37, no. 5, pp. 60–67, 2004.
[3] http://www.theguardian.com/world/2014/oct/28/ebola-australia-refuses-entry-to-people-from-liberia-sierra-leone-guinea-west-africa.
[4] M. Newman, *Networks: An Introduction*. New York, NY, USA: Oxford University Press, 2010.
[5] J. Xu, A. Kumar, and X. Yu, "On the fundamental tradeoffs between routing table size and network diameter in peer-to-peer networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 1, pp. 151–163, 2006.
[6] J. M. Kleinberg, "Navigation in a small world," *Nature*, vol. 406, no. 6798, p. 845, 2000.
[7] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," *SIGCOMM Computer Communication Review*, vol. 31, no. 4, pp. 149–160, 2001.
[8] M. Marathe and A. K. S. Vullikanti, "Computational epidemiology," *Communications of the ACM*, vol. 56, no. 7, pp. 88–96, 2013.
[9] A. Kumar, S. Merugu, J. Xu, and X. Yu, "Ulysses: a robust, low-diameter, low-latency peer-to-peer network," in *Proceedings of IEEE ICNP 2003*, pp. 258–267.
[10] https://www.dropbox.com/s/8mzxrkdefonxq86/app.pdf?dl=0.
[11] K. Savla, G. Como, M. Dahleh *et al.*, "Robust network routing under cascading failures," *IEEE Transactions on Network Science and Engineering*, vol. 1, no. 1, pp. 53–66, 2014.
[12] P. J. Lederer and R. S. Nambimadom, "Airline network design," *Operations Research*, vol. 46, no. 6, pp. 785–804, 1998.
[13] N. Wieberneit, "Service network design for freight transportation: a review," *OR spectrum*, vol. 30, no. 1, pp. 77–112, 2008.
[14] A. Ganesh, L. Massoulié, and D. Towsley, "The effect of network topology on the spread of epidemics," in *Proceedings of IEEE INFOCOM 2005*, pp. 1455–1466.
[15] F. D. Sahneh, C. Scoglio, and P. Van Mieghem, "Generalized epidemic mean-field model for spreading processes over multilayer complex networks," *IEEE/ACM Transactions on Networking*, vol. 21, no. 5, pp. 1609–1620, 2013.
[16] V. M. Preciado, F. D. Sahneh, and C. Scoglio, "A convex framework for optimal investment on disease awareness in social networks," in *Proceedings of IEEE GlobalSIP 2013*, pp. 851–854.
[17] C. Nowzari, V. M. Preciado, and G. J. Pappas, "Stability analysis of generalized epidemic models over directed networks," in *Proceedings of IEEE CDC 2014*, pp. 6197–6202.
[18] C. Enyioha, "A convex framework for epidemic control in networks," 2014.
[19] V. M. Preciado and M. Zargham, "Traffic optimization to control epidemic outbreaks in metapopulation models," in *Proceedings of IEEE GlobalSIP 2013*, pp. 847–850.

[20] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
[21] M. Caesar, M. Castro, E. B. Nightingale, G. O'Shea, and A. Rowstron, "Virtual ring routing: network routing inspired by dhts," *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 4, pp. 351–362, 2006.
[22] S. C. Ergen and P. Varaiya, "Energy efficient routing with delay guarantee for sensor networks," *Wireless Networks*, vol. 13, no. 5, pp. 679–690, 2007.
[23] S. Kniesburges, A. Koutsopoulos, and C. Scheideler, "A self-stabilization process for small-world networks," in *Proceedings of IEEE IPDPS 2012*, pp. 1261–1271.
[24] P. Van Mieghem, "Epidemic phase transition of the sis type in networks," *Europhysics Letters*, vol. 97, no. 4, p. 48004, 2012.
[25] R. K. Plowright, P. Foley, H. E. Field, A. P. Dobson, J. E. Foley, P. Eby, and P. Daszak, "Urban habituation, ecological connectivity and epidemic dampening: the emergence of hendra virus from flying foxes," *Proceedings of the Royal Society B: Biological Sciences*, vol. 278, no. 1725, pp. 3703–3712, 2011.
[26] R. Pastor-Satorras and A. Vespignani, "Immunization of complex networks," *Physical Review E*, vol. 65, no. 3, p. 036104, 2002.
[27] M. Barthélemy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, "Velocity and hierarchical spread of epidemic outbreaks in scale-free networks," *Physical Review Letters*, vol. 92, no. 17, p. 178701, 2004.
[28] H. Rahmandad and J. Sterman, "Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models," *Management Science*, vol. 54, no. 5, pp. 998–1014, 2008.
[29] https://www.dropbox.com/s/b64r3o1c0gi3iyb/codes.zip?dl=0.
[30] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, and S. Lu, "Dcell: a scalable and fault-tolerant network structure for data centers," *SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 75–86, 2008.
[31] http://openflights.org/data.html.
[32] http://apps.who.int/gho/data/view.ebola-sitrep.ebola-summary-latest?lang=en.



**Huanyang Zheng** received his B.Eng. degree in Telecommunication Engineering from Beijing University of Posts and Telecommunications, China, in 2012. He is currently a Ph.D. candidate in the Department of Computer and Information Sciences, Temple University, USA. His research focuses on wireless and mobile networks, social networks and structures, and cloud systems.

**Jie Wu** is the Associate Vice Provost for International Affairs at Temple University. He also serves as the Chair and Laura H. Carnell professor in the Department of Computer and Information Sciences. Prior to joining Tempe University, he was a program director at the National Science Foundation and was a distinguished professor at Florida Atlantic University. His current research interests include mobile computing and wireless networks, routing protocols, cloud and green computing, network trust and security, and social network applications. Dr. Wu regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including IEEE Transactions on Service Computing and the Journal of Parallel and Distributed Computing. Dr. Wu was general co-chair/chair for IEEE MASS 2006, IEEE IPDPS 2008, IEEE ICDCS 2013, and ACM MobiHoc 2014, as well as program co-chair for IEEE INFOCOM 2011 and CCF CNCC 2013. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). Dr. Wu is a CCF Distinguished Speaker and a Fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.