Identity Routing Symmetry Metrics for Routing Behavior

Fang Dong¹, a, Junbiao Liu², b and Shenyi Dai³, c

¹ School of Information & Electrical Engineering, Zhejiang University City College, China
² Department of Information Science and Electrical Engineering, Zhejiang University, China
³ College of Economics and Management, China Jiliang University, China

a dongf@zucc.edu.cn, bliujunbiao@zju.edu.cn, c daishenyi@cjlu.edu.cn

Keywords: Routing symmetry, Routing behavior model, Statistical process.

Abstract. Routing affects several protocols and impacts traffic analysis techniques. In this paper, an identity routing symmetry (IRS) metric is proposed to express network routing and its formulation is derived, which the forward and reverse flows coming from one node to another are exactly the same. IRS in three typical networks is simulated, such as random, WS and BA network. Simulation results show that routing algorithms that generate next hop randomly will lead to a symmetric network, but scale-free network does not work in this way, because degrees of nodes are varied significantly.

Introduction

In communication networks, if streams between two endpoints follow the same physical paths for both forward and reverse direction, they are symmetric. Routing asymmetry affects several protocols and impacts traffic analysis techniques [1]. Knowing to which degree the routings are symmetric is helpful in protocol design and traffic analysis.

- The one-way propagation time is commonly estimated to be half of the round-trip time between nodes, e.g., network time protocol of Internet [2]. However, this estimate will be inappropriate if routes are asymmetric. The real condition of the outgoing link in situations of routing asymmetry will determine available communication resources, thereby affecting communication protocols.

- A part of traffic analysis techniques are embedded in an assumption that routings are symmetric, i.e., all the packets of a session on both directions can be monitored by a sniffer located on a specific link. But it is not the case in practice [3]. Routing asymmetry has a significant impact of these techniques.

A common cause of routing asymmetry is that routing is selected independently for each flow. And each node takes many factors into account, including load balancing and congestion controlling. This phenomenon is especially prominent in the case of multipath routing. Another reason is the “hot-potato routing”, which is a business practice of passing traffic off to another autonomous system as soon as possible [4]. Recently, there is a growing research of macroscopic properties of the network routing, including routing asymmetry, by studying datasets or modeling network behavior [5].

In this paper, an identity routing symmetry (IRS) metric whose forward and reverse flows coming from one node to another are exactly the same is proposed to evaluate network routing. Section II lists variable symbols and models the routing selection as a Markov Chain. Section III defines IRS and derives its expression. Section IV simulates IRS performance. Section V concludes our work.

System Modeling

Notation. In this paper, matrices are denoted by boldface, uppercase letters, e.g., matrix $A$, while vectors are denoted by boldface, lowercase letters, e.g., vector $a$. The entry in the $i$-th row and the $j$-th column of matrix $A$ is referred to as $[A]_{ij}$. Identity matrix is denoted by $I$. Let $A$ and $B$ be $m \times n$ matrices. The element-wise product of $A$ and $B$ is defined by $[AB]_{ij} = [A]_{ij} [B]_{ij}$, for all $1 \leq i \leq m$, $1 \leq j \leq n$.

The variable symbols are listed in Table 1.
The routing selection as a Markov Chain. The routing probability from node \( i \) to \( j \) is the probability that node \( i \) selects a neighboring node \( j \) as the next hop to route data packets. Fig. 1 gives a further demonstration of routing probability. Fig. 1(a) shows a network with 4 nodes. Packets in Fig. 1(b) are destined to node 1. Each routing probability is labeled on a directed link from a source node to a destination node. When data packets are destined to node 4, the routing probabilities in Fig. 1(c) are different from (b). Fig. 1(d) lists the routing matrices of (b) and (c). The probability of data packet routed from node \( i \) to \( j \) in one hop is denoted by \( \rho_{ij} \). The routing probability matrix, or routing matrix for short, is given by using \( \rho_{ij} \) as the \( i \)-th row and \( j \)-th column element.

![Fig.1 Explanation of routing matrix](image-url)

There is an assumption here that routing probability is time-invariant. Our model is identical to the random walk model, which is a well-known routing behavior model. From Fig. 1, Routing probability is independent of the source node, but depends on the destination node, thereby some notes of

1. Node \( d \) is a destination node, and thus \( d \) is an absorbing state in Markov Chain, so \( [P_d]_{dd} = 1 \), and \( \sum_{j=1}^{N} [P_d]_{ij} = 1 \).
2. For any \( i \), \( \sum_{j=1}^{N} [P_d]_{ij} = 1 \).

Identity Routing Symmetry

In a connected network, there are an infinite number of paths (when counting circles) from any node \( s \) to any other node \( d \), \( path_1, path_2, path_3 \ldots \) Suppose the frequency that the source node \( s \) selects these paths to route packets to be \( p_1, p_2, p_3 \ldots \). Similarly, there are also an infinite number of paths from node \( d \) to \( s \), \( path_1^{-1}, path_2^{-1}, path_3^{-1} \ldots \) with \( path_i^{-1} \) we mean the reverse path of \( path_i \), and the corresponding selecting frequency \( q_1, q_2, q_3, \ldots \) Informally, we will use \( p \) to denote the vector \( [p_1, p_2, p_3 \ldots] \) and \( q = [q_1, q_2, q_3, \ldots] \).

### Definition of IRS metrics

We use the normalized inner product of \( p \) and \( q \) to define IRS:

\[
\rho^{id}(s,d) = \frac{(p,q)}{\sqrt{(p,p)}\sqrt{(q,q)}}, \quad (1)
\]

In algebra, \( \rho^{id} \) is also viewed as the cosine of the angle between \( p \) and \( q \). It is varied in the range \([0, 1]\). When \( \rho^{id} \) is close to 0, \( p \) and \( q \) are orthogonal to each other. So, if for some \( path_i \), the frequency that it is selected as the forward path, namely \( p_i \), is large, then the frequency that \( path_i^{-1} \) is selected as the reverse path, \( q_i \), must be small, otherwise \( \rho^{id} \) will not be close to 0. Conversely, when \( \rho^{id} \) is close to 1, \( p \) and \( q \) are parallel to one another, so for each \( i \), \( p_i \) and \( q_i \) are both large or both small. This makes \( \rho^{id} \) a good choice for defining IRS. Note that the defined \( \rho^{id} \) is not a linear function of the angle.
between \( p \) and \( q \), so we may use 1-\( \arccos(\rho^{id}) \) to calculate the identity symmetry. As this is an increasing function of \( \rho^{id} \), they are essentially the same metric.

In practical networks, usually a small number of paths are used to transfer packet stream. In such cases, \( p \) and \( q \) are sparse vectors, namely, most elements of \( p \) and \( q \) are 0s.

**Analysis of IRS.** Let nodes \( s \) and \( d \) be any two different nodes of a connected network. There is only one possible path from \( s \) to \( d \) in the network with length 1, which is \( s, d \). This path will be chosen to transfer packets with probability \([P_{d}]_{s,d}\) (if path \( s,d \) does not exist, \([P_{d}]_{s,d}\) will be zero). There are \((N-2)\) possible paths in the network with length 2, chosen with probability \([P_{d}]_{s,h_{1}}[P_{d}]_{h_{1},d}, h_{1}=1,2,\ldots,N\) and \( h_{1} \neq s, d \). There are \((N-2)^2\) paths in the network with length 2, chosen with probability \([P_{d}]_{s,h_{1}}[P_{d}]_{h_{1},h_{2}}[P_{d}]_{h_{2},d}, h_{1}=1,2,\ldots,N \) and \( h_{1}, h_{2} \neq s, d \). And the rest can be deduced by analogy. So the inner product of previously mentioned vectors \( p \) and \( q \) is

\[
(p, q) = [P_{d}]_{s,d} \cdot [P_{d}]_{d,s} + \sum_{h_{1} \neq s,d} [P_{d}]_{s,h_{1}} [P_{d}]_{h_{1},d} \cdot [P_{d}]_{d,h_{1}} [P_{d}]_{h_{1},s} + \sum_{h_{1},h_{2} \neq s,d} [P_{d}]_{s,h_{1}} [P_{d}]_{h_{1},h_{2}} [P_{d}]_{h_{2},d} \cdot [P_{d}]_{d,h_{1}} [P_{d}]_{h_{1},h_{2}} [P_{d}]_{h_{2},s} + \cdots \tag{2}
\]

All \( p_{ij} \)'s are not larger than 1 and the sum of all \( q_{ij} \)'s is 1, so the partial sum of the Right-Hand Side (RHS) of Eq.2 is less than 1, and the sequence of partial sums are incremental. So the RHS of Eq.2 is convergent. Use the definition of element-wise product operator \( \odot \), equation Eq.2 can be written in another form

\[
(p, q) = \left[ \sum_{i=1}^{\infty} (P_{d}^{(s)}) \odot (P_{d}^{(d)})^T \right]_{s,d} \tag{3}
\]

Matrix \( P_{d}^{(s)} \) is \( P_{d} \) with \( s \)-th column replaced by a zero vector, and similarly, \( P_{d}^{(d)} \) is \( P_{d} \) with \( d \)-th column replaced by zero vector.

The result of proposition in the Appendix makes it possible to write Eq.3 with a closed form. The only requirement is \( I - P_{d}^{(s)} \odot P_{d}^{(d)} \) to be invertible. Suppose this requirement is satisfied, then we get

\[
(p, q) = \lim_{n \to \infty} \left[ \sum_{i=1}^{\infty} (I - P_{d}^{(s)} \odot (P_{d}^{(d)})^T)^{-1} (P_{d}^{(s)} \odot (P_{d}^{(d)})^T - (P_{d}^{(s)} \odot (P_{d}^{(d)})^T)^{n+1}) \right]_{s,d} \tag{4}
\]

Because the \( d \)-th column and \( d \)-th row of \( (P_{d}^{(s)} \odot (P_{d}^{(d)})^T) \) are all zeros, so for any \( n \), the term \( (P_{d}^{(s)} \odot (P_{d}^{(d)})^T)^{n+1} \) will have \( d \)-th column be zeros. So,

\[
(p, q) = \left[ (I - P_{d}^{(s)} \odot (P_{d}^{(d)})^T)^{-1} (P_{d}^{(s)} \odot (P_{d}^{(d)})^T) \right]_{s,d} \tag{5}
\]

With some similar but simpler steps, we get

\[
(p, p) = \left[ \sum_{a=0}^{\infty} (P_{d}^{(s)} \odot P_{d}^{(d)})^a \cdot (P_{d} \odot P_{d}) \right]_{s,d} = \left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{s,d} \tag{6}
\]

\[
(q, q) = \left[ \sum_{a=0}^{\infty} (P_{d}^{(s)} \odot P_{d}^{(d)})^a \cdot (P_{d} \odot P_{d}) \right]_{s,d} = \left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{s,d} \tag{7}
\]

Finally,

\[
\rho^{id} = \frac{\left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{s,d} \left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{d,s}}{\sqrt{\left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{s,d} \left[ (I - P_{d}^{(s)} \odot P_{d}^{(d)})^{-1} (P_{d} \odot P_{d}) \right]_{d,s}}} \tag{8}
\]

**Simulation Results**

According to their topology, networks are usually classified into random networks, regular networks, small world (WS) networks and scale free (BA) networks. To avoid bias introduced by topology, three typical networks, namely, random, WS and BA network are considered. In the simulation, each of these 3 networks is composed of 128 nodes, thus there will be \((128 \times 127)/2 = 8128\) different pairs of nodes to be considered when evaluating identity symmetry.

Three different random walk based routing algorithms are evaluated. The routing probability from node \( i \) to \( j \) of these routing algorithms can be wrote in a unified form

\[
p_{ij} = \frac{d_{ij}}{\sum_{k \in N(i)} d_{ik}} \tag{9}
\]
but taking different values of parameter $\alpha$, namely, -1, 0, and 1, respectively. Notation $d_j$ is the degree of node $j$, and $N(i)$ the set of all neighboring nodes of node $i$.

Simulation results are shown in Fig.2. Most node pairs’ IRS are close to 1 in random network and WS network, regardless of which routing algorithm is applied. This result is consistent with assumptions that most literatures made, namely, networks are symmetric because nodes are of similar degrees in these networks, so the ratio will be approximately 1. But in scale-free network, degrees of nodes are varied significantly, thus the IRS metrics are scattered.

**Conclusion**

In this work, we propose an IRS metric to express network routing, which the forward and reverse flows coming from one node to another are exactly the same. The simulation results show that routing algorithms that generate next hop randomly will lead to a symmetric network, but scale-free network does not work in this way, because degrees of nodes are varied significantly.

**Appendix**

Proposition: For any square matrix $P$ that satisfies $I-P$ are invertible, then $\sum_{i=1}^{n} P^i = (I - P)^{-1}(P - P^{n+1})$.

**Proof:** Let $A$ be the left hand side of the equation, namely, $A = \sum_{i=1}^{n} P^i$, then $PA = P\sum_{i=1}^{n} P^i = \sum_{i=1}^{n+1} P^i$.

Subtract this equation with the above one:

$$PA - A = (I - P)A = P - P^{n+1}$$

Thus,

$$A = (I - P)^{-1}(P - P^{n+1}).$$

(10)

**Acknowledgements**

This work was financially supported by the Zhejiang Provincial Natural Science Foundation of China (LQ13G010005), Beixiu Youth Project of Zhejiang University City College (2014), and 8th key course construction project (practice) in Zhejiang University City College.

**References**


