DIT and Beyond: Inter-domain Routing with Intra-domain Awareness for IIoT

Peizhuang Cong, Yuchao Zhang, *Member, IEEE*, Lei Wang, Wendong Wang, *Member, IEEE*, Xiangyang Gong, Tong Yang, *Member, IEEE*, Dan Li, *Senior Member, IEEE*, and Ke Xu, *Senior Member, IEEE*

***Abstract*—Along with the ever-increasing amount of data generated from industrial devices, cross domain (also known as Autonomous Systems, AS) data transmission problem has attracted more and more attention in Industrial Internet of Things (IIoT). As mature and widely used inter-domain routing protocols, BGP-based solutions often take the number of domains (i.e., AS hops) of each path as a criterion to make routing decisions, which is simple and effective. However, such protocols can only meet reachability requirements while ignoring performance requirements. That is, the path with the minimum AS hops will be selected to carry flows, even if the actual performance of this path does not meet the transmission requirements due to the unawareness of intra-domain information on that path. But it is not impractical to directly access intra-domain information for making better routing decisions given data privacy concerns.**

In this paper, we propose M-DIT, which can make inter-domain routing decisions with the assistance of desensitized intra-domain information for multiple-requirement transmissions. To do so, we design a homomorphic encrypted-based private number comparison scheme to export intra-domain information securely and thus assist in routing decisions. The results of some experiments based on 5 real topologies (*ATMnet*, *Claranet*, *Compuserve*, *NSFnet*, and *Peer1*) with thousands of inter-domain flows demonstrate that M-DIT reduced flow completion time by about 60% or selected high bandwidth paths flexibly for inter-domain routing for IIoT scenarios.

***Index Terms*—inter-domain routing, transmission protocol, private number comparison**

I. Introduction

The Industrial Internet of Things (IIoT), as a vital infrastructure, facilitates the development and implementation of industrial technologies. In various fields, such as manufacturing, transportation, agriculture, energy, power grid, massive data and messages generated from IIoT [1]. For example, as shown in Figure 1, in the intelligent manufacturing scenario, the monitoring cameras upload recording files to the remote cloud server for analyzing and storing, and the industrial robots receive remote control signals from the remote cloud server [2], [3]. Generally, different services have different transmission requirements in terms of delay, bandwidth, forwarding hops, packet loss rate, etc., such as file transfer services prefer high-bandwidth routing path, while control signals transmission services require short latency. Moreover, with the decoupling of data storing and computation, such large-scale inter-domain transmission services are becoming more and more common and important [4].

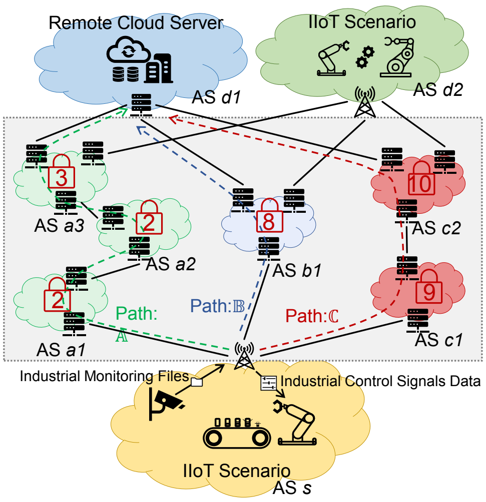


Figure 1. BGP-based inter-domain routing of IIoT scenario

As the most commonly employed inter-domain routing protocol, Border Gateway Protocol (BGP) takes the length of *AS_Path* as the routing priority metric by default [5], [6]. That is, the path with the minimum number of ASes has the highest priority [7]–[9]. Such strategy regards all domains as indiscriminate blackbox and thus cannot make performance guaranteed inter-domain routing decisions for industrial data transmission due to the lack of intra-domain information. As depicted in Figure 1, without loss of generality, assuming that an industrial terminal in AS s uploads data to a remote cloud server that belongs to AS $d1$, and there are three inter-domain paths between s and $d1$: path $\mathbb{A}$ with AS length of 4 ($s \rightarrow a1 \rightarrow a2 \rightarrow a3 \rightarrow d1$), path $\mathbb{B}$ with AS length of 2 ($s \rightarrow b1 \rightarrow d1$), and path $\mathbb{C}$ with AS length of 3 ($s \rightarrow c1 \rightarrow c2 \rightarrow d1$). Assuming the value shown in each AS represents the cost (e.g., delay) generated by crossing it, then $\mathbb{A}$ (with cost 7) is with lower accumulated cost than $\mathbb{B}$ (with cost 8) and $\mathbb{C}$ (with cost 19). However, in line with the BGP routing principle

The work was supported in part by the National Key R&D Program of China under Grant 2019YFB1802603, the National Natural Science Foundation of China (NSFC) under Grant 62172054, the NSFC under Grant 62072047, and Key Project of Beijing Natural Science Foundation under M21030.

Peizhuang Cong, Lei Wang, Wendong Wang, and Xiangyang Gong are with State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications (BUPT). Yuchao Zhang is with BUPT, Beijing, China. Tong Yang is with Peking University, Beijing, China. Dan Li and Ke Xu are with Tsinghua University, Beijing, China. Corresponding author: Yuchao Zhang (yczhang@bupt.edu.cn).

(regardless of manually specified routing rules), $\mathbb{B}$, which has the fewest AS-hops, will be selected as the forwarding path. However, $\mathbb{A}$ that outperforms $\mathbb{B}$ under the given metric will be omitted from the routing table. Therefore, it can be observed that some intra-domain information which can be leveraged to optimize inter-domain routing policies should not be ignored.

Several studies are proposed to enhance the inter-domain transmission performance by optimizing routing policies [10]–[12]. [13]–[16] employ software defined networks architecture or assign reliable service systems to compute and distribute routing policies in centralized fashions. However, the centralized fashion has two downsides: 1) it relies on specific intra-domain information to generate routing policies, which limits the deployment scope, e.g., such approaches are only suitable for the case where all domains are affiliated with trusted organizations; 2) it also suffers from unsatisfactory scalability. These downsides prevent the centralized fashion from utilizing the intra-domain information for inter-domain routing, and hence the BGP-based distributed protocol remains a practical approach. Given this, would it be practical to directly embed specific intra-domain information into the header of BGP notification message packet and diffuse it to other ASes? Such a strawman way is not practicable as ASes affiliated with different organizations may refuse to provide the required intra-domain information on account of privacy issues. Hence, bridging the gap between data sharing and privacy protecting remains a challenge.

To this end, we propose a BGP-based intra-domain state aware multi-requirement inter-domain routing policy for IIoT (M-DIT), which can be either implemented as a complement to the BGP internal functions or as a control plane function. M-DIT enables the accessibility of intra-domain information while protecting the privacy of specific data (e.g., intra-domain topology, links' status) at the same time, thus bridging the gap between data sharing and privacy protection. More specifically, M-DIT aims to represent the performance evaluation (forwarding hops, latency, bandwidth, etc.) of inter-domain routing paths; however, as stated above, it is not secure to directly share the intra-domain information by the BGP notification messages. Hence, for each metric of routing path evaluations, M-DIT basically adopts three schemes (*abstraction*, *confusion*, and *comparison*) to guarantee data privacy when notifying and diffusing intra-domain information which can facilitate inter-domain routing decision-making. M-DIT only maintains the border routers (nodes) while ignoring the specific intra-domain network topology, and builds weighted virtual connections (edges) between each pair of nodes (*Topology Abstraction*). By doing this, it not only can mask the topology and employed protocol of intra-domain but also preserve the required intra-domain information for inter-domain routing (§IV-A-1). To prevent intra-domain information (the state of links between pairs of nodes of the domain) from being leaked during route notification and diffusion, M-DIT adds a random number to each route before notifying it to neighboring domains from the source domain (*Random Number Confusion*). It protects the state privacy of the intra-domain path from the border router to the destination and does not affect the result of the routes priorities calculation (§IV-A-2). Moreover, avoid leakage of intra-domain information during route diffusing, we designed a homomorphic encryption-based algorithm that can compare priorities of paths without exposing specific values (*Private Number Comparison*, §IV-C). Further to this, M-DIT is extended to multi-requirement transmission scenarios which can provide flexible inter-domain routing decisions for different types of flows with multiple specified metrics.

We exhibit the advantages of M-DIT by embedding it in the worldwide implemented BGP using five real-world topologies and thousands of simulated flows. The results show that, for the selected representative metrics (Flow Completion Time (FCT), path bandwidth), M-DIT can enable BGP to reduce about 60% FCT on average or select high-bandwidth path preferentially for routing in multiple requirements transmission scenario. In summary, the following outlines our contributions in this paper:

- We expose that a series of BGP-based protocols unable to provide optimal inter-domain forwarding path for the multiple requirements transmission owing to the unawareness of intra-domain state.
- We propose M-DIT, which can select the optimal inter-domain path for IIoT and beyond by leveraging homomorphic encryption algorithms to sense intra-domain information without leaking it.
- We exhibit the promotions of M-DIT in contrast to traditional BGP-based protocols, by deploying some experiments on different scales in five real network topologies with multiple routing requirements.

The paper is structured as follows. We review background and related works in §II. In §III, we specify the motivation and design principle of M-DIT. In §IV, we describe M-DIT in detail. We demonstrate experimental results in §V. Lastly, we summarize this work in §VI.

II. Background and Related Work

In this part, we first present the background of this work from three aspects, i.e., BGP, multiple requirements routing, and homomorphic encryption. And, correspondingly, we exhibit developments and research status of them.

A. Background

1) Border Gateway Protocol:

Currently, as one of the most widely employed routing protocols among domains, BGP enables to glue a vast volume of ASes distributed all around the world together. Each domain takes its border gateway/router which implemented the external BGP as the egress and ingress for exchanging route entries to peers, the others inside routers execute the internal BGP. The content of the *AS_PATH* filed of each route entry indicates the length of the forwarding path in AS granularity, but it exclusively ignores the varying internal transmission capabilities of each AS.

BGP mainly includes four type messages, `OPEN`, `UPDATE`, `NOTIFICATION`, and `KEEPALIVE`. The `UPDATE` is utilized to notify and withdraw route entries, which mainly includes three features for route selecting: *AS_PATH*, *MULTI_EXIT_DISC* (*MED*) and *LOCAL_PREF*. *AS_PATH* is used to keep track of which ASes a route has crossed during transmission. The router will reject all route entries that contain its own AS number, which can be used for loop-proof and also for path selection, i.e., the shorter the *AS_PATH* the better. *MED* is announced by neighbor AS to discriminate its multiple export ports. By default, for the same neighbor AS, the lower *MED*, the higher the priority of the export port. *LOCAL_PREF* is usually configured manually by the local administrator. When an AS has multiple egress routers, the router with the largest *LOCAL_PREF* value will be set as the egress.

Manually configuring on the basis of experience is a preferred manner of inter-domain routing in current network. However, it still has limitations in inflexibility or incorrect configuration, for example, the global service disruption at Meta due to careless configurations by engineers [17]. Auto-configuration for the evolving network is becoming a developing trend. The inability of sensing the inter-domain state, the current BGP can only provide a connectivity guarantee that selects the forwarding path according to *AS_PATH* by default. However, with the ever-increasing volume of network traffic and the multiple requirements of services, the disadvantages introduced by ignoring intra-domain capabilities will be increasingly visible.

2) Multiple Requirements Routing:

Multiple requirements transmission, also referred to as multiple optimality criteria routing and multiple objective routing in some work, is a routing strategy to support the development of diverse network services. There are different priorities for different services regarding latency, bandwidth, packet loss, forwarding hops, and other metrics. Routing strategies based only on reliability or a single metric are overstretched for modern networks, which makes the research of multiple requirements transmission more valuable.

Typically, the metrics can be divided into two categories, accumulative type and bottleneck type. For accumulative type metrics, the final routing path performance is impacted by cumulative qualities of every traversed link, which is commonly calculated by addition or multiplication, such as delay, forwarding hops, packet loss rate. The final path quality corresponding to the bottleneck type metrics is determined by the extreme values of all the links traversed, which can be calculated by $min()$ or $max()$, e.g., the link bandwidth.

3) Homomorphic Encryption:

Homomorphic encryption (HE) is a cryptographic method, which can perform arithmetic calculations on ciphertext and get a equal result with encrypted form to performing specified calculations on plaintext of these ciphertext [18]. Hence, it may provide a potential solution to bridging the gap between information sharing and privacy protecting. Concretely, HE can be demonstrated as follow:

$$De(En(a) \odot En(b)) = a \oplus b, \quad (1)$$

where $En()$ is the encryption operation, $De()$ is the decryption operation, and $\odot$ and $\oplus$ are correspond to the operations on the plaintext and cyphertext domains, respectively. When $\oplus$ represents addition, this encryption is an additive homomorphic encryption, and when $\oplus$ represents multiplication, this encryption is a multiplicative homomorphic encryption. The encryption function that satisfies both additive and multiplicative homomorphism properties and can perform any times of additive or multiplicative operations is called fully homomorphic encryption.

HE algorithms, especially complete ones, suffer from high computational complexity. Nevertheless, it is merely necessary to calculate little numbers on the condition of additive HE in this work. Then, it will avoid the potential issues introduced by the complex calculation of HE algorithms. Inspired by the feature of HE algorithms, we exploit an additive HE-assisted intra-domain state sensing scheme without data leaking, whose details will be introduced in §IV.

B. Related Work

1) Enhancement of BGP: BGP is the most widely deployed inter-domain routing protocols on the Internet. There are several works dedicated to optimizing it in terms of convergence, security, etc [6], [19], [20]. M. Milani et al. aim to accelerate the BGP convergence process by decreasing the route notification time according to the domain-level topology and validate this scheme through a series of experiments [21]. J. Brenes et al. relieve the traffic losing during reconverging process of BGP by ordering the prefixes based on the unbalanced traffic distribution [8]. Alberto et al. design a route collector&beacon-based scheme to facilitate the time synchronizing between source and destination device systems of BGP route [22]. Given the achievements of blockchain in information security areas, many studies have used it to enhance the security of BGP [23]–[25]. He et al. propose a decentralized architecture based on blockchain technology, ROA chain, which specifies every AS enable to verify the route source and prevent prefix hijacking based on a globally-consistent and tamper-resistant database [26]. There are some works attempting to dedicate deep learning to address BGP issues regarding security, configuration, and more [27]–[29].

2) Inter-domain Routing Schemes: There are several works that focus on the optimization of inter-domain routing policies, which are commonly categorized into two types [5], [30]–[34]. The first type is built on the architecture with a dedicated third party (e.g., a controller) [13], [15], [16], [35]–[38]. Qiao et al. based on the idea of the software defined network to design a new software defined interconnections-based network architecture for the cross-domain scenario, which enables senders to define inter-domain forwarding path via a programmable interface [14]. Straightforwardly, Shahruz et al. exploit the strong performance of the Cloud server in terms of bandwidth and computing power to accelerate the computation and convergence of inter-domain protocols [16].

With the development of machine learning, many researches have applied it to network system. Reinforcement learning has recently been employed in traffic engineering decision-making scenarios. Xiaoyang et al. present a extensible RL-based framework with multiple layers to facilitate cross domain transmission performance [38]. Nevertheless, the same precondition of above schemes is employing dedicated managers to guarantee the effectiveness and impenetrability of corresponding data. The practical feasibility of such a idealized architecture is pending further discussion. Tunnel-based overlay architectures are concluded as the second type [39]–[43], whose essential approach is making it feasible to select a specified forwarding path by allowing ASes to establish tunnels between each other. In this circumstance, the overhead of each tunnel, i.e., the forwarding path, can be captured directly. However, the feature that tunnel information is not exposed to other ASes may lead to security issues, which makes it difficult to be accepted by network organizations. Furthermore, such tunnels are only notified within related domains for converging purpose of inter-domain routing, which may lead to undetectable traffic agreements to ISP. Given that, the deployment of these schemes may not be acceptable by ISPs.

3) Multiple Requirements Routing: As a classical problem, the multiple requirements routing research mainly involves intra-domain routing scenarios, which can be divided into two main categories [44], [45]: 1) The first is machine learning-based architecture. Lin et al. leverage reinforcement learning in intra-domain routing scenarios with multiple types of services to improve network transmission performance and utilization, etc [46]. Cong et al. extend the single-metric routing problem to multi-metric scenarios by using model fusion fashion [47]. 2) The second is algebra-based strategies. Sobrinho et al. design a network routing model with multi-metric and associated protocol, which tried to solve the inter-domain routing problem with multiple optimization criteria via a fully distributed approach [44]. Moreover, to address the delayed convergence problem present in this work, J. J. Garcia-Luna-Aceves et al. introduce DRIP, which is loop-free at every instant and can guarantee the convergence of feasible or optimal routing paths [48].

These works are hardly migrated to tackle inter-domain multiple requirements routing directly. Although Sobrinho's work has been extended to inter-domain routing [49], it is still premised on requiring the necessary intra-domain information, which is not in line with the purpose of this work.

4) Homomorphic Encryption: The widely used partial homomorphic encryption schemes include Benalol [50] and Paillier [18] algorithms for additive homomorphism, RSA [51], and EIGamal [52] algorithms for multiplicative homomorphism, and Goldwasser Micali [53] algorithm for bitwise homomorphism. These classical partial homomorphic encryption schemes are highly secure and computationally efficient and can guarantee data security and meet the computational efficiency requirements for eligible application scenarios. G. Craig proposed the first fully homomorphic encryption method according to the ideal lattice from a theoretical perspective, which caused a surge of research on fully homomorphic encryption in academia [54]. Subsequent work has been based on Gentry's work and is aimed at reducing computational overhead, improving computational efficiency, and taking into account security. Theoretically, the fully homomorphic encryption scheme is the best choice to protect data confidentiality without losing data availability, but the high overhead of the scheme, the computational model, and the high security make it impossible to be applied in practice. The high overhead of the fully homomorphic encryption scheme make it impossible to be applied in practice. However, scholars have since proposed somewhat homomorphic encryption [55], which is only applicable to low-order polynomial operations and allows only a limited number of homomorphic additions and multiplications on the encrypted data.

III. Motivation and Design Principle

In this section, we firstly introduce the motivation of this work. Then, on basis of the motivation, we further clarify the design principle of M-DIT.

A. Motivation

Path	AS-path	BGP Priority	State	Evaluation
$\mathbb{B}$:[*d1*, *b1*, *s*]	3	↑↑↑	selected	=(8)
$\mathbb{C}$:[*d1*, *c1*, *c2*, *s*]	4	↑↑	ignored	=(9,10)
$\mathbb{A}$:[*d1*, *a1*, *a2*, *a3*, *s*]	5	↑	ignored	=(2,2,3)

Figure 2. Routing table of AS s

The routing table of AS s in the above example is shown in Figure 2. Path $\mathbb{B}$ will be selected as the forwarding path due to the smallest *AS_PATH* value than that of path $\mathbb{A}$ and $\mathbb{C}$. When considering the cost or link quality of the ASes as described previously, i.e., evaluating the performance of end-to-end inter-domain transmission, however, $\mathbb{B}$ is actually not the optimal path. For example, when the value shown in each AS indicates the delay, then $\mathbb{A} \succ \mathbb{B} \succ \mathbb{C}$ ($sum(2,2,3) < sum(8) < sum(9,10)$); while, if the value indicates the bandwidth, then $\mathbb{C} \succ \mathbb{B} \succ \mathbb{A}$ ($min(9,10) > min(8) > min(2,2,3)$), where $\succ$ means *"better than"*.

How to simultaneously satisfy the requirements of exporting intra-domain information, privacy protection, and ensuring the correctness of routing calculation is the key to addressing this issue. Then, on top of this, it is possible to leverage the accumulated intra-domain data as an additional attribute of the local routing information base to aid in routing decisions. Although the motivated example is explanatory, it can be seen that the influence of intra-domain status on inter-domain transmission is non-negligible. In other words, intra-domain information awareness will be beneficial when making inter-domain routing.

B. Design Principle

To eliminate the conflict of information sharing versus data leakage, a scheme that is aware of but does not leak intra-domain data is necessary. Therefore, we specify these two requirements in detail.

- **Exporting Information:** The performances of inter-domain paths are corporately affected by abilities of all links that contained by traversed ASes, so it is necessary to notify such beneficial information along the path with a specified form to facilitate inter-domain routing. Such information mainly includes the performance evaluation of an intra-domain path regarding delay, bandwidth, packet loss, hops, and more.

- **Protecting Privacy:** For security reasons (e.g., it is possible to infer the detailed network topology of the domain by forwarding hops) or business reasons, the exported data by each AS should not be captured or inferred by others. This security guarantee is also a prerequisite for each AS to provide such relevant information.

Detail schemes of how M-DIT satisfies the design principles are exhibited in the § IV.

IV. M-DIT Methodology

In this section, without loss of generality, M-DIT is illustrated by using the accumulated forwarding hops as an evaluation metric. On this basis, we explain the differences in the computation of bottleneck-type metric and extend M-DIT to multiple requirements inter-domain routing. Finally, the incremental deployability, as well as the flexibility of M-DIT, are discussed.

A. M-DIT Overview

The field used to assist in routing path selection in the BGP header is *AS_Path* by default [5], which is also used for free-loop guarantee, then we introduce a new header field (*Attr*) to carry performance evaluation of the inter-domain path for M-DIT. Alternatively, the existing fields of BGP header (such as *MED*) can also be re-defined and re-used to simplify implementation.

1) Topology Abstraction:

In the current network architecture, on the one hand, the intra-domain routing policy is independent of the inter-domain routing protocol, that is, each domain forwards incoming traffic to the egress border router along a specified path determined by the employed intra-domain protocol; on the other hand, the inter-domain paths of BGP are granularized by border routers, which means that the next hop specified by the forwarding path is the border router of a domain [7]. Then, the performance evaluation of the path from ingress to egress of a domain is sufficient to be the intra-domain information which can be utilized to promote the generating of inter-domain routing polices.

It is reasonable to abstract each detail intra-domain topo into a graph which only contains all border routers. Given the connectivity within a domain, there are direct links or indirect connections between all border router (node) pairs, and these links and connections are recognized as edges in the graph. The *Abstraction* example is depicted in the Figure 3. It is acceptable to maintain these paths' performance of a domain. First, some protocols operating in domains or controllers of software-defined network architectures commonly maintain such information, e.g., the OSPF routing protocol maintains the forwarding hops of intra-domain paths. Alternatively, the complexity of additional maintenance of the required peer-to-peer path information is $O(N^2)$, where N is the number of border routers of a domain and is generally a small number. By doing so, it is possible to mask some intra-domain details while preserving essential information.

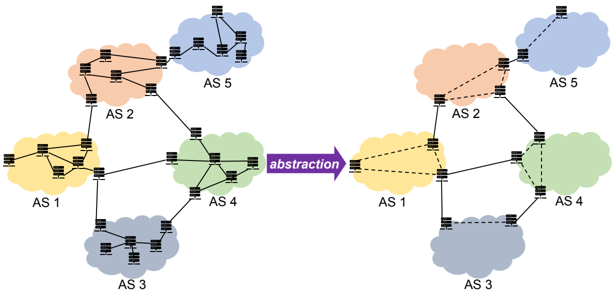


Figure 3. Topology abstraction

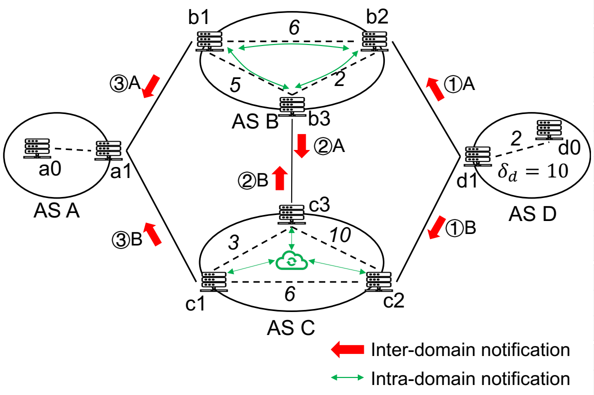
2) Random Number Confusion:

When path information (which can be assumed as the forwarding hops from the ingress to the egress router for ease of understanding) is embedded into the BGP header directly and notified to neighboring domains, the accumulated computation (addition for forwarding hops) of multiple domains during route diffusion can inherently protect the information privacy. It can be directly explained from mathematical perspective that specific values of the two elements cannot be inferred from their sum, i.e., $c \nrightarrow a$ and $c \nrightarrow b$, where $a \in \mathbb{R}$, $b \in \mathbb{R}$, and $c = a + b$. Moreover, such mathematical characteristic is one of the basic principles for privacy guarantee in M-DIT design. In the case of Figure 3, assuming that *AS 5* receives a route to a destination belonging to *AS 1* via *AS 2*, it cannot infer the specific intra-domain information corresponding to *AS 2* and *AS 1* from the cumulative path information carried in this route.

However, the inherent information protection of the aforementioned accumulated computations is only valid when such computations have been performed at least one time. For example, *AS 2* can obtain some intra-domain information about *AS 1* from the routes notified by *AS 1* that the destination belongs to it. Consequently, when notifying the route from its destination belonged domain to directly connected neighboring domains, the protection of accumulated computation will fail, which is named the *Direct Connection* ($\leftrightarrow$) leakage in this paper.

To this end, we design the *Random Number Confusion* to fix such leakage. The performance evaluation values of different paths to the same destination are only used to compare relative magnitudes, therefore the absolute values of these data do not affect the routing results as long as the relative relations remain constant. Mathematically explaining, according to the inequality principle, adding or subtracting the same value on both sides of an inequality simultaneously will not affect the comparing result. That is, if $\exists a, \exists b \in \mathbb{R}^+ \rightarrow a < b$, then $\forall c \in \mathbb{R} \rightarrow a{+}c < b{+}c$. Alternatively, it can also be understood as assigning a fixed random offset to all nodes of a coordinate system will not shift their relative positions.

M-DIT stipulates that the destination's domain adds a specified random value to the initial intra-domain information before notifying the route to neighbor domains. This process



Update Message

①A	AS D (d1)->AS B (b2)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	d1	1	12	D
①B	AS D (d1)->AS C (c2)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	d1	1	12	D
②A	AS B (b3)->AS C (c3)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	b2	3	14	D
②B	AS C (c3)->AS B (b3)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	c1	2	21	D
③A	AS B (b1)->AS A (a1)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	b2	3	19	D
③B	AS C (c1)->AS A (a1)			
Des	Next-hop	AS-path	Attr	Des-AS
d0	c3	4	17	D

RIB

b2	New Attr = Attr : maintain			
Des	Next-hop	AS-path	Attr	Des-AS
d0	d1	2	12	D
c2	New Attr = Attr : maintain			
Des	Next-hop	AS-path	Attr	Des-AS
d0	d1	2	12	D
c3	New Attr < Attr : update			
Des	Next-hop	AS-path	Attr	Des-AS
d0	c2-> b3	2-> 3	22-> 17	D
b3	New Attr > Attr : maintain			
Des	Next-hop	AS-path	Attr	Des-AS
d0	b2	2	14	D
c1	updated by intra-AS notification			
Des	Next-hop	AS-path	Attr	Des-AS
d0	c2-> c3	2-> 3	18 ->17	D
a1	New Attr < Attr : update			
Des	Next-hop	AS-path	Attr	Des-AS
d0	b1-> c1	3-> 4	18 ->17	D

Figure 4. Diffusion illustration: MSGs diffusion and RIB updates of M-DIT triggered by new routes

can be defined as:

$$d_{Notified} = d_{initial} + \delta_d, \tag{2}$$

where $d_{Notified}$, $d_{initial}$, and δ_d are the notified value, the initial value, and the corresponding specified random value of the routes with destination d, respectively. The indeterminate $delta_d$ makes it impossible for neighboring domains to obtain the corresponding intra-domain information, which can also remain the correctness of the route computation result during subsequent route diffusion.

As a result, M-DIT is able to address the *Direct Connection*($\leftrightarrow$) leakage by employing accumulated computation coupled with *Random Number Confusion* without shifting the routing selection.

3) Information Diffusion:

In this work, we define and add a new field *Attr* for BGP header to carry the mentioned data. However, this solution is optional and it is feasible to redefine and reuse existing fields, such as *MED*. Then, the quantified evaluation of the routing path performance will be written in *Attr* of the BGP update message (MSG).

Whether the domain is running traditional protocols (AS_B of Figure 4) or based on software defined architecture (AS_C), it is permitted as long as the *Attr* field can be processed accurately according to M-DIT. Assuming that the route of *d0* is updated, AS_D (*d1*) will send this update to AS_B (*b2*) and AS_C (*c2*). The *Attr* of update MSG is 12 (δ_d + 2). *b2*/*c2* determines whether updates local route or not by comparing *Attr* value of received MSG with local route. The corresponding route will be refreshed if its *Attr* is greater than the newly received *Attr*. And vice versa. In the interior of AS_B/AS_C, this route update will be exchanged by intra-domain protocol/controller. After the internal exchange, AS_B diffuses the update MSG to AS_C, where *Attr* is summed by two components: 1) the performance of intra-domain forwarding path (*b3*, *b2*)(the qualified value is 2); 2) the *Attr* value received from *b2* (12). That is, the value of *Attr* in this MSG is 14 (2+12). Similarly, AS_C(*c3*) sends a update MSG with *Attr* = 21 (3+6+12) to AS_B(*b3*). *c3* will update the corresponding local route due to the received *Attr* from *b3* is smaller than the local value. On the contrary, *b3* will do not modify local RIB. Then, AS_B (*b1*)/AS_C (*c1*) send update MSG to AS_A (*a1*) with *Attr* is 19 (*b1*, *b3*, *b2*, *d1*)/17 (*c1*, *c3*, *b3*, *b2*, *d1*). Finally, *a1* updates the forwarding path to destination *d0* for AS_A according to these received messages. Then, *a1* will assign *c1* as the next hop for traffic with destination *d0* based on the new route entries.

B. Delta Trap

The proposed schemes so far appear to guarantee data privacy, but there is still a potential risk of information leakage during routing diffusion. Then, this leakage risk will be introduced for convenience from the description of a simple mathematical problem. Given b_1 and b_2 are known, where b_1 equals $a_1 + a_2 + a_3$ and b_2 equals $a_1 + a_2$. It is possible to obtain a_3 from the difference (*Delta Trap*, Δ) between b_1 and b_2, even if both a_1 and a_2 are unknown ($a_3 = b_1 - b_2$). This problem is mapped to the information leakage problem in route diffusion as follows: in Figure 4, after receiving the routes about *d0* from AS_D and AS_C successively, due to the information of *AS_Path*, AS_B (*b3*) can obtain intra-domain information about AS_C based on the difference between the two routes' *Attr* (the intra-domain routing policy of *c2* to *c3* and corresponding path state, [*c2*$\rightarrow$*c1*$\rightarrow$*c3*]), which is a potential risk for AS_C. However, adding a random value to *Attr* is not applicable for routes with destinations outside local domains. The proposed *Random Number Confusion* would shift the result of subsequent route computation, which can be described as "$x_1 > x_2 \nrightarrow x_1 > x_2 + \delta \,|\, x_1, x_2, \delta \in \mathbb{R}$" from a mathematical perspective.

To address the aforementioned issues, we further propose *Private Number Comparison* for M-DIT.

C. Enhanced M-DIT

Delta Trap (Δ) is caused by receiving two routes destined for the same destination, where one traverses one additional domain than the other. Topologically describing, the domains in a triangular connection suffer from such information leakage risk during route diffusion. This risk can be eliminated by breaking the triangular connection in the topology provided keeping the forwarding path unaffected, i.e., logically masking links between connected domains that are not on the optimal routing path.

To this end, we propose *Private Number Comparison*, which can complete the comparison calculations without disclosing the specific values of all three parties. Then, the comparison results can assist in masking non-optimal links from the topology during route diffusion. In the following, the adopted homomorphic encryption algorithm and the specific workflows of *Private Number Comparison* will be presented in detail.

1) Homomorphic Encryption:

The cryptosystem generally uses public/private keys to encrypt/decrypt the plaintext/ciphertext. Paillier, a classical homomorphic encryption method [18], is employed in this work, whose processes of keys generation, encryption, and decryption and homomorphism of addition are as follows.

- *Key Generation*: Randomly selecting two large prime numbers p and q that satisfy $gcd(pq,(p-1)(q-1))=1$, and calculating $n=pq$ and $\lambda=lcm(p-1,q-1)$. And randomly selecting integer $g\in\mathbb{Z}^*_{n^2}$, and calculating $\mu=(L(g^\lambda \, mod\, n^2))^{-1} mod\, n$, where $L(u)=\frac{u-1}{u}$, for $\forall u\in\{u<n^2 \mid u=1\, mod\, n\}$. Then, the public key is (n,g) and private key is (λ,μ).
- *Encryption*: For plaintext $m\in\mathbb{Z}^*_n$, its encrypted ciphertext is $c=g^m\cdot r^n\, mod\, n^2$.
- *Decryption*: For ciphertext $c\in\mathbb{Z}^*_{n^2}$, its decryped plaintext is $m=L(c^\lambda\, mod\, n^2\cdot\mu)\, mod\, n$.

Assuming that $r_1,r_2\in\mathbb{Z}^*_{n^2}$ are two random integers, for the plaintext m_1, m_2, their ciphertext are $En(m_1)\equiv g^{m_1}\cdot r_1^n\, mod\, n^2$ and $En(m_2)\equiv g^{m_2}\cdot r_2^n\, mod\, n^2$, respectively. Then,

$$\begin{aligned} En(m_1)\cdot En(m_2) &\equiv g^{m_1}\cdot r_1^n\cdot g^{m_2}\cdot r_2^n\, mod\, n^2 \\ &\equiv g^{m_1+m_2}\cdot(r_1\cdot r_2)^n\, mod\, n^2 \\ &\equiv En(m_1+m_2) \end{aligned}$$

As $r_1,r_2\in\mathbb{Z}^*_{n^2}$, then $r_1\cdot r_2\in\mathbb{Z}^*_{n^2}$, so the Paillier cryptosystem is additive homomorphic. Hence, in this work, Paillier is subtly applied during the number comparison process to prevent specific values from being disclosed.

2) Private Number Comparison:

Traps Detection. It is necessary to detect triangular connections from the domain topology. The first step is adjacent domains exchange locally maintained neighboring domains list. The second step is that each domain calculates the corresponding triangle connection according to Algorithm 1. This process can be operated by specific applications or existing BGP messages.

Although this algorithm is designed for the case of directed links between domains, it can still be adapted to the undirected link scenario by simply removing duplicated triangle elements.

Algorithm 1: Δ detection

get_neighbors(AS): get AS's neighbor list
Input: neighboring domains lists
Output: the triangular connections list of AS
for i *in get_neighbors(*AS*)* **do**
  for j *in get_neighbors(*i*)* **do**
    res.append([AS, i, j])
return res

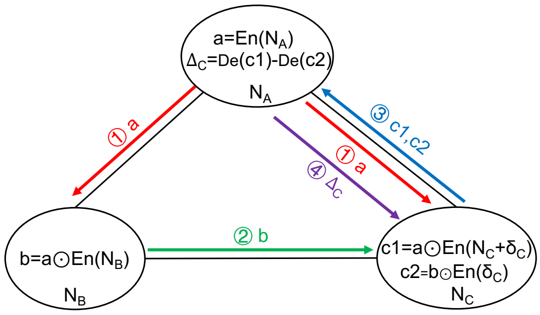


Figure 5. Workflows of *Private Number Comparison*

Moreover, the triangle connection remains stable provided that links between domains remain unchanged.

Comparing Paths. In the generic triangular topology, as shown in Figure 5, path comparison and selection would be accomplished by communicating with each other, which is described as pseudo code Algorithm 2.

Suppose *A*, *B* and *C*, each of which is responsible for local values, N_A, N_B, N_C, respectively. First, A sends encrypted N_A by private key of *A*, $En^A(N_A)$, to B and *C*. After receiving the MSG from *A*, *B* sends $En^A(N_A)\odot En^A(N_B)$ to *C*, where $\odot$ represents homomorphic addition calculation, which means $En(x)\odot En(y)\equiv En(x+y)$. After receiving the MSG from *A* and *B*, *C* sends $En^A(N_A+N_B)\odot En^A(\delta_C)$ and $En^A(N_A)\odot En^A(N_C+\delta_C)$ to *A* in the specified order. After receiving the MSG from *C*, *A* decrypts and subtracts the two values, $De^A(En^A(N_A+N_B+\delta_C))-De^A(En^A(N_A+N_C+\delta_C))$, and get the signed delta value Δ_C, which will be sent back to *C*. Finally, according to Δ_C, *C* and *A* can determine the priority of the two paths, $Path_{(C\to A)}$ and $Path_{(C\to B\to A)}$.

The reason why A has to send N_A to B and C is that the cost of C or B through A to the same border router of A during inter-domain transmission may be different, i.e., the N_A sent by A to B and C is the respective corresponding cost, and this comparison algorithm is still feasible.

The confidentiality of the entire comparison process is explained here. The value sent by A to B and C is encrypted and cannot be decrypted by B and C with public keys, and likewise, the value sent by B to C cannot be deciphered. The malicious case of forcing to break the encryption algorithm is not considered here. The two values sent by C to A use the confusion strategy by adding a random value, which makes A

Algorithm 2: Comparison

En(x): encrypt x
De(x): decrypt x
Send([x_1, x_2], [D_1, D_2]): send [x] to [D]
Rec(MSG): receive message MSG
Input: the connection of (A, B, C)
Output: comparing result
AS A:
  ena=En(N_A, key_A)
  A.Send(ena, [B, C]) `// marked MSG1`
AS B:
  na = B.Rec(MSG_1)
  enb = $na \odot$ En(N_B, key_A)
  B.Send(enb, C) `// marked MSG2`
AS C:
  na = C.Rec(MSG_1)
  nb = C.Rec(MSG_2)
  enc = $na \odot$ En($N_C + \delta_C$, key_A)
  enb = $nb \odot$ En(δ_C, key_A)
  C.Send([enc, enb], A) `// marked MSG3`
AS A:
  nc, nb = A.Rec(MSG_3)
  nb = De(nb, KEY_A)
  nc = De(nc, KEY_A)
  Δ_C = nc - nb
  A.Send(Δ_C, C) `// marked MSG4`
AS C:
  Δ_C = C.Rec(MSG_4)

only can get Δ_C after decryption.

Constraining Diffusion. In the case of Figure 4, if the AS_C does not receive the route update from AS_D, it would not cause intra-domain information leakage and will also properly update the local RIB based on the route received from AS_B. Therefore, M-DIT constrains the route diffusion for the triangle-connected domains on the basis of the comparison results. As shown in Figure 6, the constraint can be divided into two cases: a) if *B* and *C* forward traffic with the same destination via *A* as the corresponding optimal path, then *A* will set a flag "$TAG_{\triangle} = 1$" when notified of the related route to declares that *B* and *C* are forbidden to notify this route to each other; b) it is assumed without loss of generality that path [$C \to B \to A$] is better than [$C \to A$], then *A* only notifies the corresponding route to *B* with flag "$TAG_{\triangle} = 0$", which means *B* can notify this route to *C*. And *C* will inherently avoid notifying this route back to *A* according to the loop-free property of BGP.

D. Completeness Analysis of Privacy

The enhancement of inter-domain routing by leveraging intra-domain information requires protecting the data of each domain from being obtained by others, which can be modeled

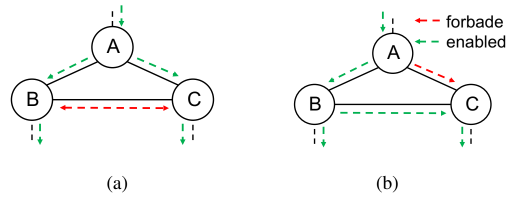


Figure 6. Constraints illustration

as an equation solving problem. During the convergence of a route, domain S accumulates maintained local transmission performance $cost_S$ (e.g., forwarding hops) on the *Attr* coming upstream and spreads downstream. Thus, each domain can obtain an equation $Attr_i = \sum_J^{AS\_PATH_i} cost_J$ based on the *Attr* and *AS_PATH* of route i. Privacy protection aims to prevent any domain from inferring any $cost_J$ from a series of equations generated by different routes of local RIB. In the following, we first mathematically model the problem and then prove the privacy completeness of the M-DIT.

1) Formulation: We define the cumulative cost for domain S of being forwarded by its border router j to domain D is:

$$COST_S^{j \to D} \tag{3}$$

Then, based on different n border routers, S can obtain the set of equations $\mathbb{C}$:

$$COST_S^{i \to D} = y_i, i \in n, \tag{4}$$

where y_i is the value of *Attr* of each related route i. Each $COST_S^{j \to D}$ can be expressed in the form of a cumulative sum of the *cost* of route i. So the set $\mathbb{C}$ can be converted as:

$$cost_{iAS_0}^D + cost_{iAS_1}^D + ... + cost^D = y_i, i \in n, \tag{5}$$

where $cost_{iAS_j}^D$ represents the cost of the j*-th* domain of the path forwarded by the border router i to domain D. Intra-domain data leakage occurs when any *cost* can be inferred from $\mathbb{C}$.

2) Mathematical Analysis: For the first case, if exists $i \in [0, n-1]$ that makes $\sum_{j=1}^{D} \text{cost}_{ij}^D$ in $\mathbb{C}$, i.e., the aforementioned *Direct Connection* ($\leftrightarrow$), it is straightforward to obtain that $cost_{i0}^D = y_0$. This situation is solved by random number confusion.

For the second case, if there is no intersection of the paths, i.e., there are no identical domains on the paths except for the end domain. At this point of $\mathbb{C}$, the number of unknowns is greater than the number of equations, so no unique solution can be derived.

For the third case, there are intersections in multiple paths, we have the following theorem.

Theorem IV.1. *If there are overlapped ASes on any two routing paths to the same destination, then the sub-paths of these two paths from the overlapped AS to the destination are the same.*

Proof. Assuming that the two sub-paths from the overlapped domain to the destination are different, i.e., there are more than one optimal paths to the destination from the overlapped

AS, which contradicts the principle that each domain will only choose one optimal path to the destination. That is, the assumption is not valid. □

Based on Theorem IV.1, we represent the domains before the overlapped domain as $\sum_{j=1}^{D} \text{cost}_{ij}^{D}$. Then, the equation corresponding to the paths with overlapped domain, $\mathbb{K}$, can be converted as:

$$cost_{l}^{D} + cost_{l^1}^{D} + ... + \sum_{j=1}^{D} \text{cost}_{ij}^{D} = y_l,\ l \in \mathbb{K} \tag{6}$$

According to the property of a system of non-homogeneous linear equations, the necessary and sufficient condition for the equation system $Ax = b$ to have a solution is that the rank of the coefficient matrix is equal to the rank of the augmented matrix, i.e., $rank\,(A) = rank(A, b)$, and the necessary and sufficient condition for having a unique solution is $rank\,(A) = n$. For Eq (6), **iff** $\exists a \in \mathbb{K} \rightarrow \sum_{j=1}^{D} \text{cost}_{ij}^{D} = y_t$, and $\exists b \in \mathbb{K} \rightarrow cost_{b}^{D} + \sum_{j=1}^{D} \text{cost}_{ij}^{D} = y_t$, it can uniquely infer the value of an unknown quantity. That is, $cost_{b}^{D} = y_t - y_b$. This situation corresponding to the aforementioned *Delta Trap* (Δ), which can be solved by private number comparison. Hence, it is capable to guarantee M-DIT 's privacy.

E. Multiple Requirements Routing

Before extending the single metric inter-domain routing scheme to multiple metrics, we describe the difference between the desensitization procedure for bottleneck type and the aforementioned desensitization procedure for cumulative type. First, the abstraction process remains consistent, i.e., masking specific topologies and states inside the domain and preserving connections between border routers. Second, in the random number confusion process, when the source domain notifies a route entry, it is required to assign an initial value to the metric evaluation. For example, for bandwidth, a relatively large value or desired bandwidth will be set so that the subsequent $min()$ calculation would not be biased. Conversely, if a smaller evaluation value for a metric is preferred, the initial value should be zero to ensure the correctness of the subsequent $max()$ calculation. Finally, since the target of the private number comparison process is to compare the priority of two paths, the process remains fundamentally consistent. The only difference is whether the path with a larger or smaller value should be specified according to the characteristics of the corresponding metric.

Based on the above, the inter-domain routing scheme can be extended to multiple requirements scenarios in two implementation ways. The key concerns of transmission are mainly concentrated on a few metrics, e.g., latency, bandwidth, packet loss, etc. Then, the first implementation employs a straightforward and efficient way of notifying routes independently for different metrics. Such a fashion not only ensures the convergence independence of each metric, i.e., only the corresponding route should be converged when a metric of paths changes, but also enables flexible adjustment of the weights of concerned metrics in routing decisions according to requirements. It decouples the various requirements of transmission services and route entries, thus maximizing the flexibility of routing decisions. The second implementation is embedding values of multiple metrics into specified several sequenced *Attr* fields the packet header. In route notification messages, the *Attr* field associated with the metric that is forbidden to diffuse or do not need to be reconverged will be filled with 0, which indicates that this *Attr* field is unavailable. The calculation of each field is executed independently following the aforementioned operations.

Assuming that each *Attr* is f bits, the notification message excluding the *Attr* field is N bits, n metrics need to be maintained in the network, and the number of messages generated by once convergence of metric$_i$ is P_i. Then the overhead generated by once convergence of the first implementation is $(N + f) * \sum_{i=1}^{n} P_i$, and that of the second implementation is $(N + nf) * max(P_i | i \in [1, n])$. It is possible to choose a more efficient way depending on the network demand following the calculation. For convenience, we employ the first implementation scheme in this work.

F. Discussion of Incremental Deployment and Flexibility

Given the scale of the existing Internet, it is impossible to deploy M-DIT all at once, although it can enhance inter-domain transmission performance. Therefore, incremental deployability is necessary. The nature of M-DIT is to desensitize intra-domain data to assist in inter-domain routing, thus such information can be carried by protocols other layers to pass through the domains that do not support M-DIT. That is, M-DIT can be converged in incremental deployment scenarios. In this case, assuming that several paths have the same length of *AS_Path*, it is possible to: 1) specify that paths with a higher proportion of non-M-DIT have lower priority; 2) specify the evaluation of the non-M-DIT domain path as the average performance of all M-DIT domains; 3) discard paths that will cross domains with poor performance directly. The above strategy may reduce the traffic crossed the non-M-DIT domain, which in turn may affect their revenue [56]. Therefore, M-DIT motivates each domain to deploy it from business and performance enhancement perspectives.

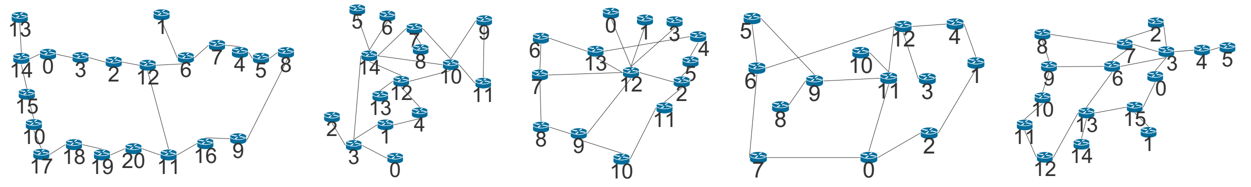
In addition, M-DIT does not interfere with each domain's behavior regarding cross-domain traffic. Firstly, M-DIT does not force each domain to specifically provide the optimal links for inter-domain traffic, instead only requires sharing the performance evaluation of links that it would provide; secondly, M-DIT allows domains to egress traffic based on existing routing algorithms, such as *hot potato* routing algorithms, or to assign an ingress for traffic by setting the best evaluation to the path from the destination to the ingress border router.

V. Evaluation

In this part, we first describe the experiment settings. Then, we comprehensively analyse M-DIT 's improvements over BGP demonstrated by a series of experiments.

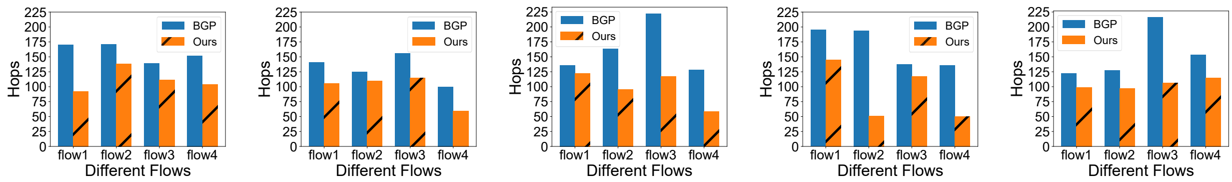
A. Experiment Setup

The network simulated by NS3 (dce-ns3-dev) on the Ubuntu 16.04.7-LTS operating system. The server is equipped with 8G



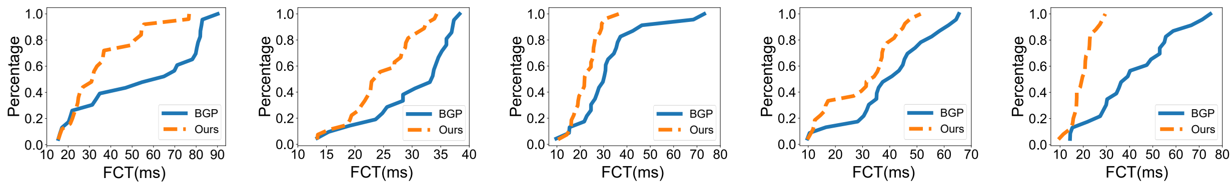
(a) ATMnet (b) Claranet (c) Compuserve (d) NSFnet (e) Peer1

Figure 7. Topology sketches of experiments



(a) ATMnet (b) Claranet (c) Compuserve (d) NSFnet (e) Peer1

Figure 8. Hops improvements



(a) ATMnet (b) Claranet (c) Compuserve (d) NSFnet (e) Peer1

Figure 9. FCT improvements

of RAM, dual core Intel(R) Core(TM) i5-6300HQ 2.30GHz CPU, and 128GB HDD. We use five real network topologies, *ATMnet*, *Claranet*, *Compuserve*, *NSFnet*, and *Peer1*, selected from *Topology Zoo* [57] for evaluation. As shown in Figure 7, each node of the topology represents an AS that implicitly contains a number of border routers and internal routers set in experiments. We set the latencies for all links with a uniform distribution of $U(0.5ms, 4.0ms)$. Moreover, we randomly select a few links and increased their latency with a probability distribution of $U(20.0ms, 50.0ms)$ to simulate the uncertain performance of links in practical networks. We generate simulated IIoT flows by referring to existing works [58], [59].

Existing inter-domain routing optimization schemes are based on centralized architectures, which derive the global view based on explicit intra-domain information. However, the prerequisite of M-DIT is guaranteeing the privacy of intra-domain information. Therefore, this section mainly demonstrates the improvement over classical BGP.

B. Experiment Results

In the following, we first evaluated the performance of M-DIT in forwarding hops metric and the corresponding improvement in delay, i.e., flow completion time. Then, we integrated the bandwidth metric to analyze the performance of multiple requirements routing. Finally, we investigated M-DIT in terms of the effects of large-scale traffic and intra-domain scale, convergence performance and the required computational overhead.

1) Performance of Forwarding Hops:

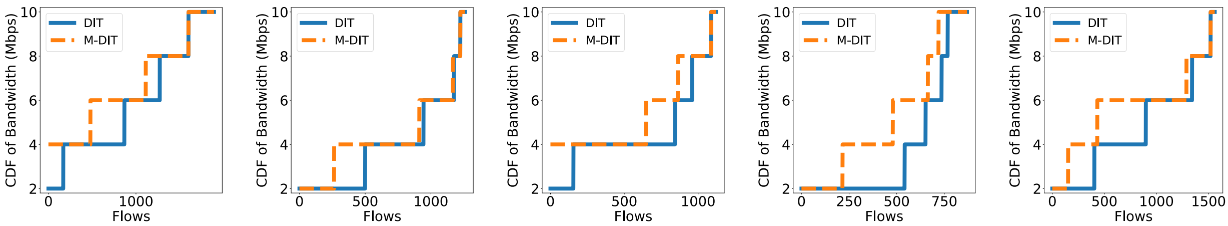
We randomly generated some flows on the selected topologies, whose source and destination are distant apart, which enables multiple paths to better demonstrate the enhancement in terms of inter-domain routing of M-DIT over BGP.

In this set of experiments, we mainly measured the metric of forwarding hops as described in §IV. In the five topologies, for flows with the same source and destination, the corresponding point-to-point forwarding hops for M-DIT and BGP are shown in Figure 8. The experimental results indicate that M-DIT can leverage the additional information to select the routing path more properly than BGP in the case with multiple inter-domain paths.

M-DIT and BGP would have the same routing policy for the flow that only has one single forwarding path, which generally exists between adjacent domains. That is, there is no room for optimization in this case for M-DIT. Therefore, it is unnecessary to measure the performance of all flows between all pair nodes, which will be present in §V-B5 in detail.

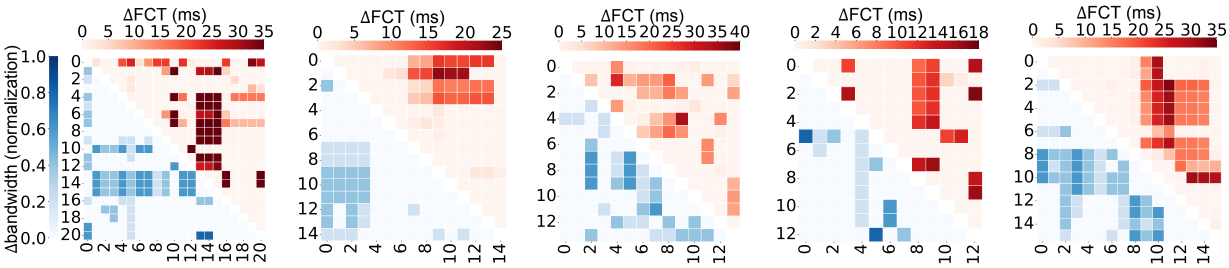
2) Performance of FCT:

Through the experiment above, we found that while taking the forwarding hops as the optimization metric, the corresponding Flow Completion Time (FCT) can also be reduced. This is because a reduction of forwarding hops can reduce the total processing delay at switches/routers for a flow. Therefore,



(a) ATMnet (b) Claranet (c) Compuserve (d) NSFnet (e) Peer1

Figure 10. M-DIT versus DIT on performance of bandwidth-sensitive flows (We added these charts of supplemented experiment analysis, and this explanation will be removed in the formal version)



(a) ATMnet (b) Claranet (c) Compuserve (d) NSFnet (e) Peer1

Figure 11. Multiple Requirements Routing Performances

it is possible to switch the FCT optimization to the more stable forwarding hops metric.

Then, specifically, we measured the FCT of flows generated in the above experiments under running M-DIT and BGP, respectively, i.e., point-to-point inter-domain transmission latency of the flows. The results shown in Figure 9 indicate that M-DIT can outperform BGP regarding FCT of generated flows on the selected topologies, despite targeting forwarding hops as the optimization metric.

3) Multiple Requirements Routing:

To illustrate the enhancement of M-DIT in a multiple requirements scenario, we generate, delay-sensitive and bandwidth-sensitive, two typical types of flows between any pair of ASes in these five network topologies, which prefer the routing path with the shortest delay and the routing path with the largest bandwidth, respectively. There is a router inside each AS that is specified to receive and send flows, which is directly connected with the border routers. In addition, to exemplify the impact of intra-domain state only, the bandwidth of all links between ASes is set to 10*Mbps*, and the bandwidth of all links inside ASes is randomly set to 10*Mbps*, 8*Mbps*, 6*Mbps*, 4*Mbps*, or 2*Mbps*. The delay of each link is kept consist as above experiments. The M-DIT is implemented in the network with both delay and bandwidth metrics.

a) Necessity: We compared M-DIT and single-metric DIT [1], where DIT takes the number of hops as the optimization goal. DIT performs the same routing policy for bandwidth-sensitive and delay-sensitive flows, both of which follow the minimum hops principle. M-DIT can leverage the multiple attributes to appropriately select inter-domain paths for both types of flows, which means that M-DIT have comparable average performance with DIT on thousands of delay-sensitive flows, as shown in the Table I.

Table I
AVERAGE FCT OF DELAY-SENSITIVE FLOWS

(# ms) - / -	*ATMnet*	*Claranet*	*Compuserve*	*NSFnet*	*Peer1*
DIT	22.258	12.471	14.137	14.758	10.489
M-DIT	22.260	12.469	14.136	14.757	10.487

Therefore, this experiment mainly demonstrates and analyzes the comparisons of bandwidth-sensitive flows performance between M-DIT and DIT. As shown in Figure 10(a) to Figure 10(e), the results indicate that M-DIT improves the transmission performance on bandwidth on average 12.94% (11.24%, 17.49%, 37.42%, and 21.24%) over DIT in *ATMnet* (*Claranet*, *Compuserve*, *NSFnet*, and *Peer1*). Although the improvement in bandwidth metrics is traded with the reduction of FCT, it is reasonable and acceptable for bandwidth-sensitive flows.

Based on the above analyses, it is beneficial and significant to provide different routing policies for different types of flows. The performances of M-DIT versus BGP will be further evaluated in the following.

b) Outperformance: Figure 11(a) to Figure 11(e) show the performance improvements of M-DIT over BGP, where the values of axes indicate the source and destination index. The undirected nature of the link properties set in the experiment makes the transmission performance symmetric for the same pair of ASes. Thus, we integrate the results of these two metrics, i.e., the upper part of the heat map indicates the ΔFCT (BGP minus M-DIT) of the selected routing paths, while the lower part indicates the Δbandwidth (normalized value of

Table II
IIoT SERVICE EMULATION PERFORMANCE

Service Type	Protocol	*ATMnet*	*Claranet*	*Compuserve*	*NSFnet*	*Peer1*
Signal Transmission Average FCT (#ms)	BGP	34.4 / 35.1 / 79.1	31.4 / 34.4 / 35.3	35.2 / 37.2 / 72.2	58.4 / 64.4 / 45.5	32.3 / 30.5 / 73.2
	M-DIT	22.1 / 11.5 / 32.2	24.0 / 25.2 / 21.2	23.0 / 17.2 / 20.1	40.3 / 11.7 / 38.2	19.2 / 20.0 / 21.2
Files Uploading Path Bandwidth (#Mbps)	BGP	2 / 2 / 6	2 / 2 / 2	2 / 4 / 2	2 / 10 / 10	4 / 2 / 4
	M-DIT	10 / 6 / 6	4 / 6 / 6	6 / 6 / 8	2 / 10 / 10	6 / 6 / 6

M-DIT minus BGP). Moreover, the darker color indicates a more significant improvement of M-DIT. The experimental results show that M-DIT can also provide routing policies that better than or at least equal to BGP in multiple requirements transmission scenario.

4) IIoT service emulation:

Referring to existing works, two typical IIoT services are involved in this experiment: monitoring files uploading with the size of 2MB per file, which represent bandwidth-sensitive services [58]; control signal transmission with flows size randomly of 30B, 50B or 100B, which represent a series of delay-sensitive services [59]. Without loss of generality, three pairs of ASes are selected as source and destination ASes for emulated services in 5 topologies (*ATMnet*: 13 to 1, 13 to 16, 1 to 9; *Claranet*: 0 to 8, 2 to 10, 9 to 2; *Compuserv*: 6 to 2, 8 to 4, 7 to 2; *MSFnet*: 8 to 3, 5 to 11, 10 to 5; *Peer1*: 14 to 5, 14 to 8, 9 to 1). Since it is only necessary to consider the service characteristics in the end-to-end inter-domain transmission, hundreds of aforementioned flows are randomly generated with equal probability at the egress node of the source AS based on [58], [59].

The average FCT of control signal transmission and the bandwidth of forwarding path of the files uploading between three pairs of selected source and destination ASes in each of five topologies are shown in Table II. The smaller the value of the FCT, the better, and vice-versa for the bandwidth. In all AS-pairs selected in this experiment, compared with BGP, M-DIT averagely reduces 49.28% FCT for delay-sensitive flows and selects a 2.03x bandwidth routing path for bandwidth-sensitive flows. The results indicate that M-DIT outperforms the BGP for inter-domain transmission for IIoT services and maintains a similar improvement with the above experiments, which further confirms the superiority of M-DIT.

5) Performance of Full-set Flows:

To completely analyze the capability of M-DIT, we simultaneously generated almost 900 different flows for all pairs of border nodes (full-set flows) in *NSFnet*. The flow completion times of such full-set flows were respectively measured under M-DIT and BGP. Since M-DIT is overall superior to BGP in terms of forwarding hops theoretically, we reasonably analyze M-DIT the performance under full-set flows in terms of latency.

With full-set flows, the result of the BGP-based flow completion time exceeds the M-DIT-based flow completion time for each flow is exhibited in Figure 12, where larger values (i.e. redder elements) indicate better performance for M-DIT. The results demonstrate that M-DIT enabled most flows to outperform or at least equalling BGP in terms of FCT in large-scale flows. It is reasonable that the performance of roughly 0.5% of flows forwarded by M-DIT is slightly inferior to BGP (negative values/light blue elements of Figure 12) due to uncertain fluctuations, e.g., packets queuing or link congestion. Overall, the result objectively indicates that M-DIT shows significant general improvements over BGP.

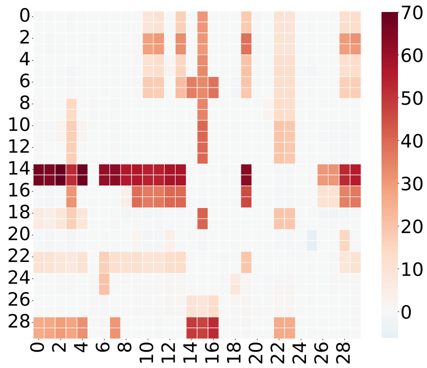


Figure 12. FCT under full-set flows

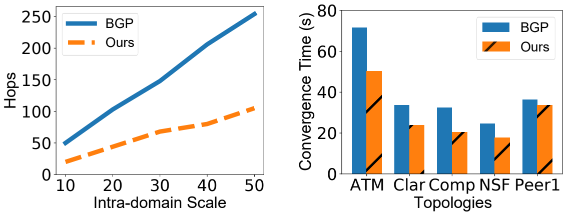
6) Influence of Intra-domain Scale:

The influence of different intra-domain scales (i.e., the number of routers within a domain) on the boost that M-DIT can achieve varies. We individually measured the forwarding hops of intra-domain scales from 10 to 50, where intra-domain links were randomly generated. Moreover, in each experiment setting, the scales of a few domains were extended beyond the assigned scale to simulate some uncertain cases.

Figure 13(a) displays the average forwarding hops for flows with different intra-domain scales under M-DIT and BGP, where the flow setting remains consistent with §V-B1. The larger the domain scale corresponds to a bigger intra-domain performance difference, and as the accumulated effect of inter-domain transmission, the final improvement in performance becomes more pronounced. The results indicate that M-DIT reduces the average point-to-point forwarding hops up to 60% for the intra-domain scale of 50.

7) Convergence and Cryptogram Overhead:

Moreover, the convergence of inter-domain protocols is an essential metric. Based on the selected five topologies, we compared the convergence times of M-DIT and BGP. The result of Figure 13(b) indicates that M-DIT outperforms classic BGP in terms of convergence, which is because M-DIT pre-pruned some route diffusion path, thus speeding up the convergence process. Additionally, the M-DIT is an independent process that precedes the route diffusion, so additional computa-



(a) Intra-domain scales (b) Convergence time

Figure 13. Protocol convergence and influence of intra-domain scale

tions (e.g., homomorphic encryption/decryption/computation, comparison) will not affect the convergence process of the protocol.

In the experiments, we also evaluate the computation overhead associated with homomorphic encryption. We precompute sufficient available primes for selection to promote computational efficiency, and stipulate that the three computations of encrypting, homomorphic addition, and decrypting are specified as an operation. The results show that the *Paillier*, implemented by Python (NTL library of C language [60]), incurs a time overhead of $30ms$ ($0.1ms$) per operation that averaged over 10^5 computations.

VI. Conclusion

In this work, we demonstrate the potential and benefit of intra-domain state awareness for multiple requirements inter-domain routing. However, it is not well-supported by existing inter-domain protocols for privacy reasons in IIoT scenarios and beyond. Given all this, we design an intra-domain state-aware inter-domain routing scheme that can securely leverage intra-domain information to enhance inter-domain routing decisions. Specifically, we exploit homomorphic encryption algorithms to secure intra-domain information, thus avoiding potential private data leaking induced by information sharing. The experimental results on five real network topologies exhibit that our proposed scheme outperforms the existing BGP-based protocols. M-DIT reduced FCT by about 60% or selected high bandwidth paths flexibly for inter-domain routing in IIoT scenarios and beyond.

Acknowledgement

This paper extends [1] by adding the solution of inter-domain multiple requirements routing for IIoT scenario and beyond.

References

[1] P. Cong, Y. Zhang, L. Wang, H. Ni, W. Wang, X. Gong, T. Yang, D. Li, and K. Xu, "Break the blackbox! desensitize intra-domain information for inter-domain routing," in *2022 IEEE/ACM 30th International Symposium on Quality of Service (IWQoS)*. IEEE, 2022, pp. 1–10.

[2] S. R. Pokhrel, L. Pan, N. Kumar, R. Doss, and H. L. Vu, "Multipath tcp meets transfer learning: A novel edge-based learning for industrial iot," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10 299–10 307, 2021.

[3] M. A. Rahman, M. S. Hossain, A. J. Showail, N. A. Alrajeh, and A. Ghoneim, "Ai-enabled iiot for live smart city event monitoring," *IEEE Internet of Things Journal*, 2021.

[4] U. Cisco, "Cisco annual internet report (2018-2023) white paper," *www.cisco.com*, 2020.

[5] B. Schlinker, H. Kim, T. Cui, E. Katz-Bassett, H. V. Madhyastha, I. Cunha, J. Quinn, S. Hasan, P. Lapukhov, and H. Zeng, "Engineering egress with edge fabric: Steering oceans of content to the world," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017, pp. 418–431.

[6] B. Al-Musawi, P. Branch, and G. Armitage, "Bgp anomaly detection techniques: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 377–396, 2017.

[7] Y. Rekhter, T. Li, and S. Hares, "Rfc 4271: A border gateway protocol 4 (bgp-4)," *[Online] https://datatracker.ietf.org/doc/html/rfc4271*, 2006.

[8] J. Brenes, A. García-Martínez, M. Bagnulo, A. Lutu, and C. Pelsser, "Power prefixes prioritization for smarter bgp reconvergence," *IEEE/ACM Transactions on Networking*, vol. 28, no. 3, pp. 1074–1087, 2020.

[9] R. Fezeu and Z. L. Zhang, "Anomalous model-driven-telemetry network-stream bgp detection," in *2020 IEEE 28th International Conference on Network Protocols (ICNP)*, 2020.

[10] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu *et al.*, "B4: Experience with a globally-deployed software defined wan," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 3–14, 2013.

[11] Z. Yang, Y. Cui, B. Li, Y. Liu, and Y. Xu, "Software-defined wide area network (sd-wan): Architecture, advances and opportunities," in *2019 28th International Conference on Computer Communication and Networks (ICCCN)*, 2019.

[12] P. Lin, J. Bi, S. Wolff, Y. Wang, A. Xu, Z. Chen, H. Hu, and Y. Lin, "A west-east bridge based sdn inter-domain testbed," *Communications Magazine IEEE*, vol. 53, no. 2, pp. 190–197, 2015.

[13] K. Lakshminarayanan, I. Stoica, S. Shenker, and J. Rexford, *Routing as a Service*. Citeseer, 2004.

[14] Q. Xiang, J. Zhang, K. Gao, Y.-s. Lim, F. Le, G. Li, and Y. R. Yang, "Toward optimal software-defined interdomain routing," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. IEEE, 2020, pp. 1529–1538.

[15] G. Asharov, D. Demmler, M. Schapira, T. Schneider, G. Segev, S. Shenker, and M. Zohner, "Privacy-preserving interdomain routing at internet scale." *Proc. Priv. Enhancing Technol.*, vol. 2017, p. 147, 2017.

[16] S. Pouryousef, L. Gao, and A. Venkataramani, "Towards logically centralized interdomain routing," in *17th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2020, pp. 739–757.

[17] B. Barrett, "Why facebook, instagram, and whatsapp all went down today," *Wired*, 2020.

[18] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *International conference on the theory and applications of cryptographic techniques*. Springer, 1999, pp. 223–238.

[19] Q. Li, J. Liu, Y.-C. Hu, M. Xu, and J. Wu, "Bgp with bgpsec: Attacks and countermeasures," *IEEE Network*, vol. 33, no. 4, pp. 194–200, 2018.

[20] M. Chiesa, A. Kamisiński, J. Rak, G. Rétvári, and S. Schmid, "A survey of fast-recovery mechanisms in packet-switched networks," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1253–1301, 2021.

[21] M. Milani, M. Nesler, M. Segata, L. Baldesi, L. Maccari, and R. L. Cigno, "Improving bgp convergence with fed4fire+ experiments," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2020, pp. 816–823.

[22] A. García-Martínez and M. Bagnulo, "Measuring bgp route propagation times," *IEEE Communications Letters*, vol. 23, no. 12, pp. 2432–2436, 2019.

[23] M. Saad, A. Anwar, A. Ahmad, H. Alasmary, M. Yuksel, and A. Mohaisen, "Routechain: Towards blockchain-based secure and efficient bgp routing," in *2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. IEEE, 2019, pp. 210–218.

[24] S. Angieri, M. Bagnulo, A. García-Martínez, B. Liu, and X. Wei, "Inblock4: Blockchain-based route origin validation," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2020, pp. 291–296.

[25] D. Chen, Y. Ba, H. Qiu, J. Zhu, and Q. Wang, "Isrchain: Achieving efficient interdomain secure routing with blockchain," *Computers & Electrical Engineering*, vol. 83, p. 106584, 2020.

[26] G. He, W. Su, S. Gao, and J. Yue, "Securing route origin authorization with blockchain for inter-domain routing," in *2020 IFIP Networking Conference (Networking)*. IEEE, 2020, pp. 504–508.

[27] M. Bahnasy, F. Li, S. Xiao, and X. Cheng, "Deepbgp: a machine learning approach for bgp configuration synthesis," in *Proceedings of the Workshop on Network Meets AI & ML*, 2020, pp. 48–55.

[28] K. McGlynn, H. Acharya, and M. Kwon, "Detecting bgp route anomalies with deep learning," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2019, pp. 1039–1040.

[29] A. Lutu, M. Bagnulo, C. Pelsser, O. Maennel, and J. Cid-Sueiro, "The bgp visibility toolkit: Detecting anomalous internet routing behavior," *IEEE/ACM Transactions on Networking*, vol. 24, no. 2, pp. 1237–1250, 2015.

[30] A. Gupta, R. MacDavid, R. Birkner, M. Canini, N. Feamster, J. Rexford, and L. Vanbever, "An industrial-scale software defined internet exchange point," in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2016, pp. 1–14.

[31] R. R. Sambasivan, D. Tran-Lam, A. Akella, and P. Steenkiste, "Bootstrapping evolvability for inter-domain routing with d-bgp," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017, pp. 474–487.

[32] J. L. Sobrinho, D. Fialho, and P. Mateus, "Stabilizing bgp through distributed elimination of recurrent routing loops," in *2017 IEEE 25th International Conference on Network Protocols (ICNP)*. IEEE, 2017, pp. 1–10.

[33] T. Holterbach, S. Vissicchio, A. Dainotti, and L. Vanbever, "Swift: Predictive fast reroute," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017, pp. 460–473.

[34] K.-K. Yap, M. Motiwala, J. Rahe, S. Padgett, M. Holliman, G. Baldus, M. Hines, T. Kim, A. Narayanan, A. Jain *et al.*, "Taking the edge off with espresso: Scale, reliability and programmability for global internet peering," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 2017, pp. 432–445.

[35] A. Gupta, L. Vanbever, M. Shahbaz, S. P. Donovan, B. Schlinker, N. Feamster, J. Rexford, S. Shenker, R. Clark, and E. Katz-Bassett, "Sdx: A software defined internet exchange," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 551–562, 2014.

[36] R. Birkner, A. Gupta, N. Feamster, and L. Vanbever, "Sdx-based flexibility or internet correctness? pick two!" in *Proceedings of the Symposium on SDN Research*, 2017, pp. 1–7.

[37] A. Dethise, M. Chiesa, and M. Canini, "Prelude: Ensuring inter-domain loop-freedom iñ sdn-enabled networks," in *Proceedings of the 2nd Asia-Pacific Workshop on Networking*, 2018, pp. 50–56.

[38] X. Zhao, C. Wu, and F. Le, "Improving inter-domain routing through multi-agent reinforcement learning," in *IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2020, pp. 1129–1134.

[39] W. Xu and J. Rexford, "Miro: Multi-path interdomain routing," in *Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications*, 2006, pp. 171–182.

[40] X. Yang, D. Clark, and A. W. Berger, "Nira: a new inter-domain routing architecture," *IEEE/ACM transactions on networking*, vol. 15, no. 4, pp. 775–788, 2007.

[41] H. Wang, Y. R. Yang, P. H. Liu, J. Wang, A. Gerber, and A. Greenberg, "Reliability as an interdomain service," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 229–240, 2007.

[42] S. Peter, U. Javed, Q. Zhang, D. Woos, T. Anderson, and A. Krishnamurthy, "One tunnel is (often) enough," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 99–110, 2014.

[43] Y. Wang, J. Bi, and K. Zhang, "A sdn-based framework for fine-grained inter-domain routing diversity," *Mobile Networks and Applications*, vol. 22, no. 5, pp. 906–917, 2017.

[44] J. L. Sobrinho and M. A. Ferreira, "Routing on multiple optimality criteria," in *Proceedings of the Annual conference of the ACM Special Interest Group on Data Communication on the applications, technologies, architectures, and protocols for computer communication*, 2020, pp. 211–225.

[45] P. Cong, Y. Zhang, W. Wang, and K. Xu, "Soho-fl: A fast reconvergent intra-domain routing scheme using federated learning," *IEEE Network*, 2023.

[46] C. Liu, M. Xu, Y. Yang, and N. Geng, "Drl-or: Deep reinforcement learning-based online routing for multi-type service requirements," in *IEEE INFOCOM 2021-IEEE Conference on Computer Communications*. IEEE, 2021, pp. 1–10.

[47] P. Cong, Y. Zhang, Z. Liu, T. Baker, H. Tawfik, W. Wang, K. Xu, R. Li, and F. Li, "A deep reinforcement learning-based multi-optimality routing scheme for dynamic iot networks," *Computer Networks*, vol. 192, p. 108057, 2021.

[48] J. Garcia-Luna-Aceves, B. R. Smith, and J. T. Samson, "Qos routing using dominant-distance vectors," in *2022 IEEE/ACM 30th International Symposium on Quality of Service (IWQoS)*. IEEE, 2022, pp. 1–10.

[49] J. L. Sobrinho and M. A. Ferreira, "From non-optimal routing protocols to routing on multiple optimality criteria," *IEEE/ACM Transactions on Networking*, 2022.

[50] J. D. C. Benaloh, "Verifiable secret-ballot elections," Ph.D. dissertation, Yale University, 1987.

[51] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120–126, 1978.

[52] T. ElGamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE transactions on information theory*, vol. 31, no. 4, pp. 469–472, 1985.

[53] S. Goldwasser and S. Micali, "Probabilistic encryption," *Journal of computer and system sciences*, vol. 28, no. 2, pp. 270–299, 1984.

[54] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the forty-first annual ACM symposium on Theory of computing*, 2009, pp. 169–178.

[55] J. Fan and F. Vercauteren, "Somewhat practical fully homomorphic encryption." *IACR Cryptol. ePrint Arch.*, vol. 2012, p. 144, 2012.

[56] D. Gupta, A. Segal, A. Panda, G. Segev, M. Schapira, J. Feigenbaum, J. Rexford, and S. Shenker, "A new approach to interdomain routing based on secure multi-party computation," in *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 2012, pp. 37–42.

[57] Zoo, "The internet topology zoo," http://www.topology-zoo.org/, 2021.

[58] T. Hussain, K. Muhammad, J. Del Ser, S. W. Baik, and V. H. C. de Albuquerque, "Intelligent embedded vision for summarization of multiview videos in iiot," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2592–2602, 2019.

[59] X. Wang, H. Yao, T. Mai, Z. Xiong, F. Wang, and Y. Liu, "Joint routing and scheduling with cyclic queuing and forwarding for time-sensitive networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 3, pp. 3793–3804, 2022.

[60] V. Shoup *et al.*, "Ntl: A library for doing number theory," 2001.

Peizhuang Cong received the Ph.D. degree from the State Key Laboratory of Network and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, China. His current research interests include next generation network architecture, data-driven networks, network resources management and QoS, and mobile internet.

Yuchao Zhang (Member, IEEE) received the B.S. degree in computer science and technology from Jilin University in 2012 and the Ph.D. degree from the Department of Computer Science, Tsinghua University, in 2017. She is currently an Associate Professor with the Beijing University of Posts and Telecommunications, Beijing, China, and a Visiting Scholar with the University of Cambridge, where she is also a Research Associate at the Wolfson College. Her research interests include large scale datacenter networks, federated learning, data-driven networks, and edge computing. She is a member of ACM.