

网络配置智能化探索与实践

李福亮 2023年11月30日

目 录

1 配置综合概述

2 配置综合实践

3 配置翻译

4 配置大模型





Part 01

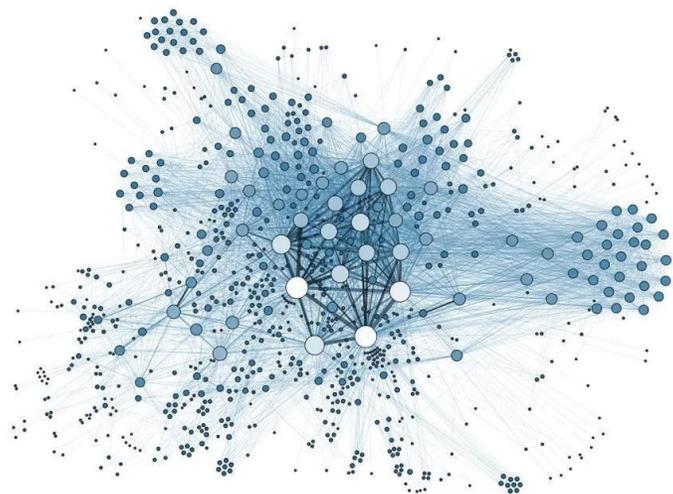
配置综合概述



复杂网络配置管理

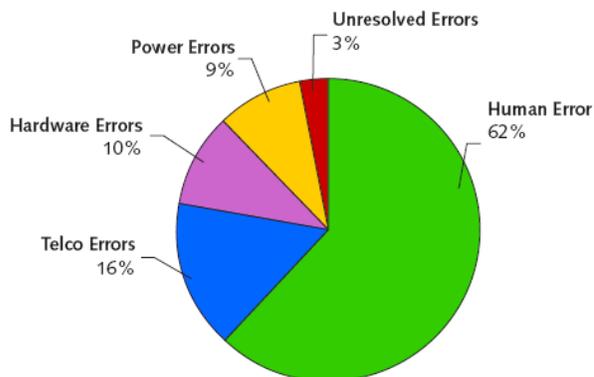
网络越来越复杂，网络设备支持的服务和功能越来越多，导致手动配置错误经常出现。

- ✓ 62%的网络中断事件与**配置错误**有关 —— Yankee Group
- ✓ **配置错误**导致的网络故障接近70% —— 阿里巴巴
- ✓ 97%的人认为人为**配置错误**会导致网络故障的发生 —— Veriflow



Human Error Is the Leading Cause of Network Downtime

Source: The Yankee Group 2002 Network Downtime Survey



ISP 配置 BGP 错误导致谷歌云瘫痪，中国电信背了黑锅。。。

云头条 11月14日

近日众多美国用户连接至谷歌（包括谷歌云、YouTube及其他网站）的流量突然经过俄罗斯重新路由至中国境内。

这就意味着，得克萨斯州、加利福尼亚州和俄亥俄州等州的网民启动浏览器和软件，并连接到谷歌及其服务后，不是与属于这家硅谷巨头的服务器进行联系，而是与俄罗斯和中国的系统进行联系。美国境外的网民也可能受到了影响。

谷歌已证实，在星期一下午太平洋标准时13:12至14:35的一段时间，连接至谷歌云、API及网站的流量经过属于境外互联网服务提供商（ISP）的IP地址改道而行。在谷歌云上构建的网站和应用程序（比如Spotify、Nest和Snapchat）也因这次拦截而中招。

在BGP路由问题导致全球性故障持续六个多小时后，Facebook、Instagram和WhatsApp开始重新上线。



今天美国东部标准时间上午11点50分前后，这三大网站都突然无法访问，浏览器在尝试打开它们时显示DNS错误。

配置参数繁多

配置命令种类及其参数繁多是导致错误的重要原因之一。下面仅是Cisco一个型号下的配置命令种类

IP Routing

- Cisco IOS IP Routing: BGP Command Reference
- Cisco IOS IP Routing: EIGRP Command Reference
- Cisco IOS IP Routing: ISIS Command Reference
- Cisco IOS IP Routing: LISP Command Reference
- Cisco IOS IP Routing: ODR Command Reference
- Cisco IOS IP Routing: OSPF Command Reference
- Cisco IOS IP Routing: Protocol-Independent Command Reference
- Cisco IOS IP Routing: RIP Command Reference
- Cisco IOS Performance Routing Command Reference

基本配置命令超1000余种

包括IP组播、路由、LAN交换、MPLS等功能
IP转发功能中，仅BGP相关命令文档有800余页



命令包含大量参数

各个参数有各自含义和限制

命令、命令参数间相互关联



BGP的一条命令

redistribute (IP)

To redistribute routes from one routing domain into another routing domain, use the **redistribute** command in the appropriate configuration mode. To disable all or some part of the redistribution (depending on the protocol), use the **no** form of this command. See the “Usage Guidelines” section for detailed, protocol-specific behaviors.

```
redistribute protocol [process-id] {level-1 | level-1-2 | level-2} [autonomous-system-number]
[metric {metric-value | transparent}] [metric-type type-value] [match {internal |
external 1 | external 2}] [tag tag-value] [route-map map-tag] [registered] [resolved]
[summary] [subnets] [nssa-only]
no redistribute protocol [process-id] {level-1 | level-1-2 | level-2} [autonomous-system-number]
[metric {metric-value | transparent}] [metric-type type-value] [match {internal | external 1 |
external 2}] [tag tag-value] [route-map map-tag] [registered] [resolved] [summary]
[subnets] [nssa-only]
```

配置参数填充难

填充这些命令、参数不仅要了解其含义，还要清楚与其他配置之间的关联。

比如：

- 接口cost设置要符合厂商要求，还要注意全局转发行为的影响
- OSPF可以用network，也可以import来引入，如import要考虑相关协议通告
- 修改community-list时，要考虑应用它的route-map等网络特性的影响

```
interface TenGigabitEthernet1/1/1
  ip address ??
  ip ospf cost 10 < ? < 100

router ospf 100
  ?
  ...

router bgp 6500
  ...
  neighbor AS200 import route-map imp-p1
  neighbor AS200 export route-map exp-p1
  ...
ip community-list C1 permit ?
ip community-list C2 permit ?

route-map imp-p1 permit 10
  ?
route-map exp-p1 ? 10
  match community C2
route-map exp-p2 ? 20
  match community C1
  ...
```

一种解



```
interface TenGigabitEthernet1/1/1
  ip address 10.0.0.1 255.255.255.254
  ip ospf cost 15

router ospf 100
  network 10.0.0.1 0.0.0.1 area 0.0.0.0

router bgp 6500
  ...
  neighbor AS200 import route-map imp-p1
  neighbor AS200 export route-map exp-p1
  ...
ip community-list C1 permit 6500:1
ip community-list C2 permit 6500:2

route-map imp-p1 permit 10
  set community 6500:1
  set local-pref 50
route-map exp-p1 permit 10
  match community C2
route-map exp-p2 deny 20
  match community C1
  ...
```

配置参数关联性强

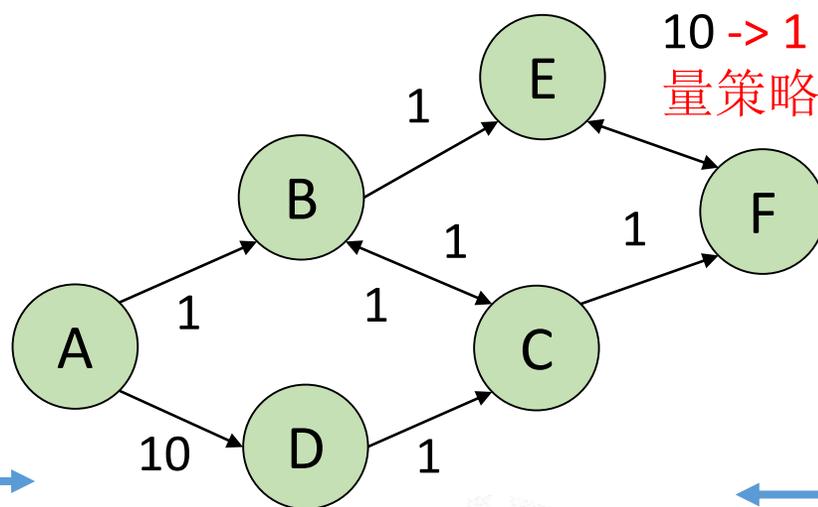
网络配置是一件复杂、需要很多专业技能与经验的工作。因为网络配置的低层次、分布式以及各种协议自身的原因导致关键参数与语句配置困难。

OSPF协议cost做流量路径规划

流量路径需求

x.x.x.10 : A,B,C,F

x.x.x.20 : D,C,B,E



10 -> 1 .10和.20流量策略将不满足

新增流量路径需求

x.x.x.30 : E,F

根据需求设置cost需要全局考量

修改一个参数可能会对全局流量产生影响

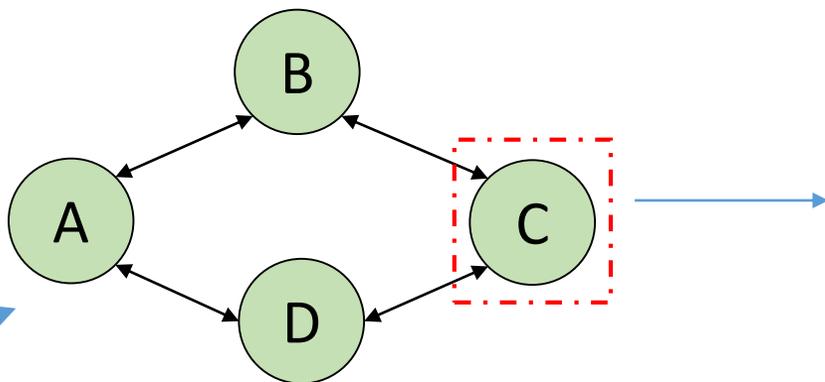
配置导致安全隐患

ACL规则数量巨大，单设备级/全网级都存在冗余失效的问题

安全需求



配置到具体设备



流量的转发路径

需要分析转发表或配置得到

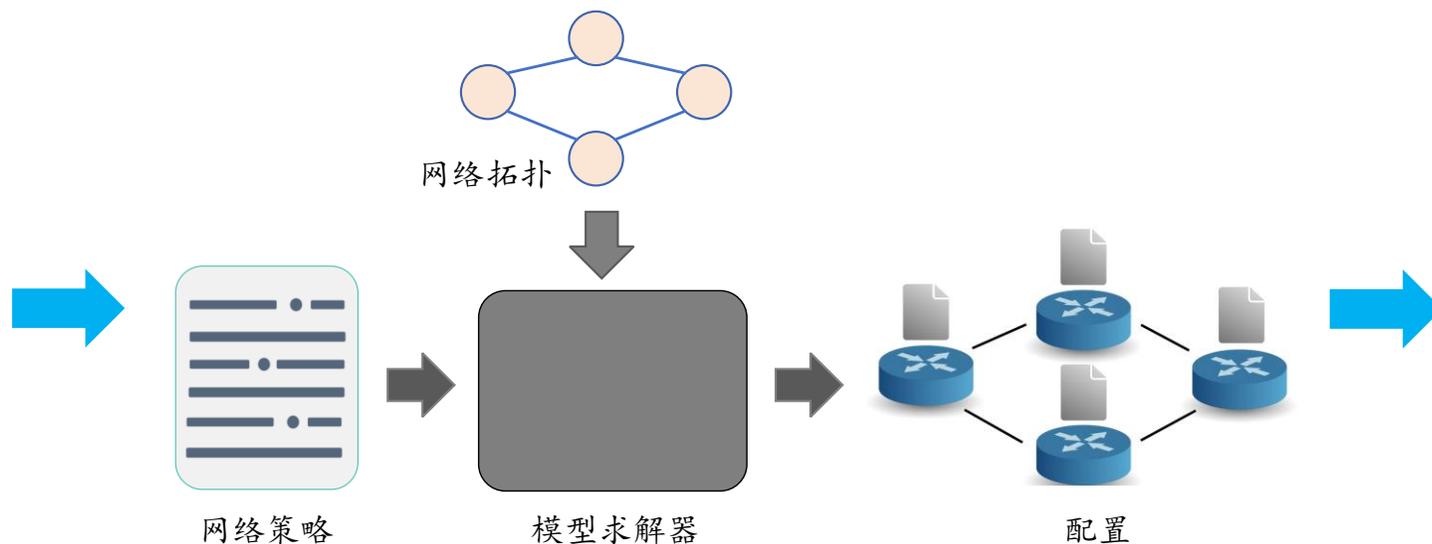
若ACL没配在对应的流量路径上则修改规则无效，导致安全隐患或配置冗余。

单个设备上ACL规则数量非常庞大，且可能存在规则action混用的情况。管理员为使当前安全需求生效往往会将新规则最先生效，进一步加大规则的混乱，使ACL规则存在大量失效、冗余的语句，加大了安全隐患。

复杂网络配置管理

配置错误原因:

- ✓ 配置参数繁多
- ✓ 配置命令低级
- ✓ 网络设备多样性
- ✓ 网络服务需求多
- ✓ 配置关联性较强
- ✓ 人工手动的配置

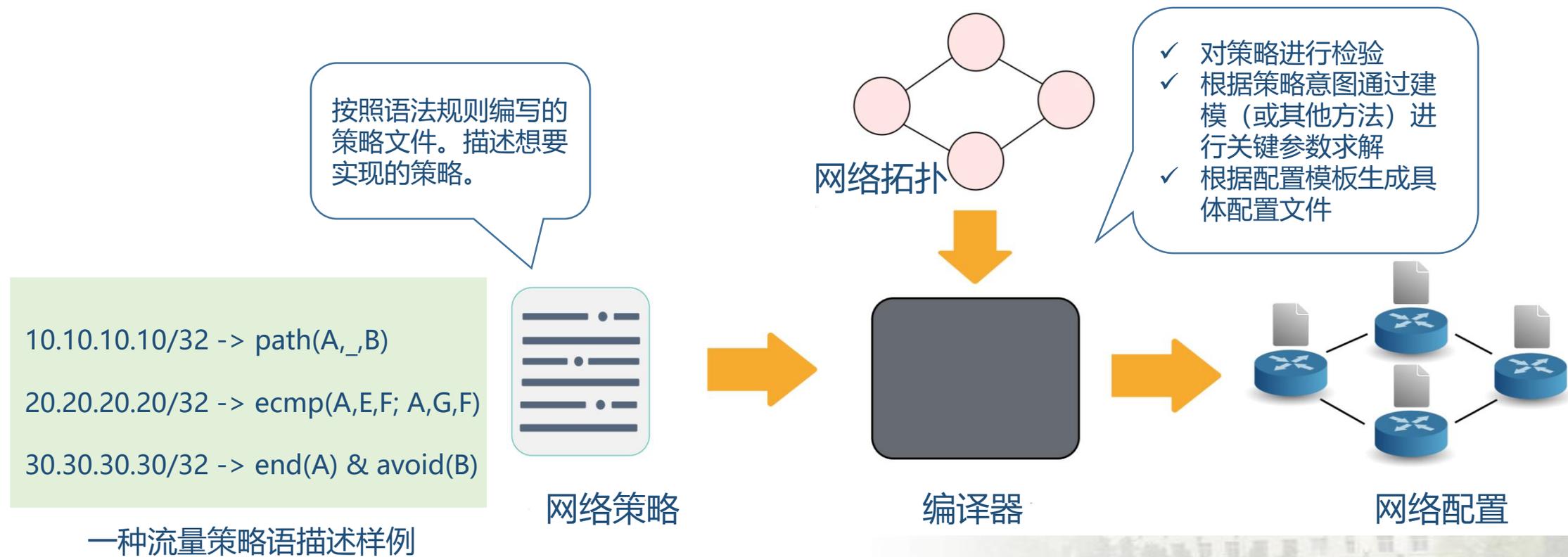


网络配置自动生成模型

自动生成复杂的网络配置，并自动化部署，减少人为错误带来的网络异常和潜在风险漏洞。

配置综合技术 (Configuration Synthesis)

- ✓ 声明式语言表达配置意图，描述目标网络想要到达的状态。
- ✓ 编译器检验用户配置策略并综合生成具体配置文件。



一种流量策略语描述样例

网络配置DSL

领域特定语言 (Domain-Specific Language)

- ✓ 为某一领域实现特定功能而设计的程序语言
- ✓ SQL语言是数据库的DSL
- ✓ **网络配置DSL**: 用以实现网络配置自动生成的程序语言

结构逻辑性差

策略描述模糊

编译处理复杂

优点: 通俗易懂

“目的IP为10.10.10.10/32
的流量从节点C流出网络,
并且不能经过节点D”

“目的IP为20.20.20.0/24
的流量有两条负载均衡路径
为A,B,C和A,D,C”

“目的IP为30.30.30.30/32
的流量在路径B,C,D出故障
时走备选路径B,E,D”

自然语言

对比

10.10.10.10/32 -> exit(C) & avoid(D)

20.20.20.0/24 -> ecmp(A,B,C ; A,D,C)

30.30.30.30/32 -> order(B,C,D ; B,E,D)

领域特定语言

结构逻辑性强

策略表述清晰

编译处理简单

缺点: 语法学习

配置综合验证 (DEBUG)

就像程序需要DEBUG一样，为保证配置综合的正确输出，要对用户输入的策略进行验证。

✓ 语法错误

用户编写的策略文件可能存在语法错误

```
10.10.10.10/32 -> ecmpp(A,B,D ; A,C,D)
```

关键字拼写错误

```
20.20.20.20/32 -> end(C)  
20.20.20.20/32 -> end(B)
```

同一流量重复描述

```
10.10.10.10/32 -> ecmp (A,nD ; A,C,D)
```

拓扑中不存在的节点

✓ 语义错误

用户编写的策略文件可能存在冲突或无法实现

```
20.20.20.20/32 -> ospf(A,C,D)  
30.30.30.30/32 -> ospf(A,B,D)
```

OSPF协议无法同时实现这两条最优路径

```
10.10.10.10/32 -> path (A,B,C,B,D)
```

最优路径为环路

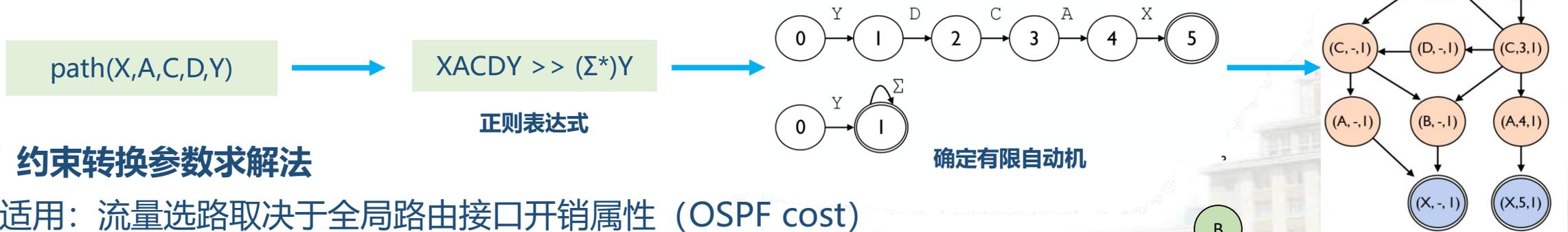
配置综合实现

根据策略类型（流量，安全），协议类型（BGP, OSPF, SR, ACL）和实现功能（生成，更新）设计相应综合方法，计算相关参数，自动生成具体配置。

流量策略的两种实现方法：

✓ 状态机生成路径图法

适用：流量选路取决于路由节点本地优先级属性（BGP local-preference）



✓ 约束转换参数求解法

适用：流量选路取决于全局路由接口开销属性（OSPF cost）



BGP综合实现样例

✓ 状态机生成路径图法

根据语法规则，将转发需求转换为正则表达式。同一流量的多条路径约束分别进行表示， \gg 表示其优先关系。

正则表示结合拓扑生成确定有限自动机，该自动机抽象表示了满足转发需求的流量状态变化的逆过程。

对BGP协议进行建模，模拟路由宣告过程，匹配自动机生成带状态的路径图，根据状态变化信息生成路由策略参数。

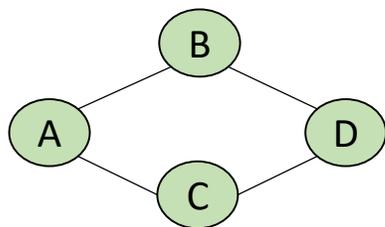
从A到B的流量走A,C,D,B
优先于走A,B

转发需求

语法规则

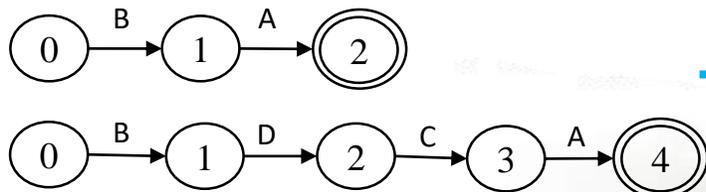
ACDB \gg AB

正则表达式

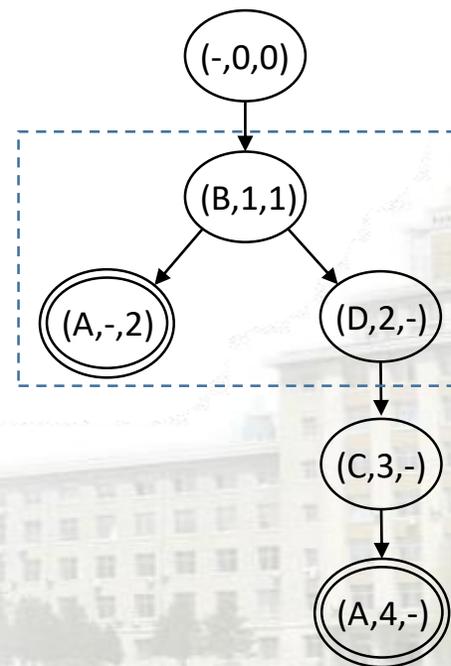


拓扑

补充输入



确定有限自动机



在B配置路由策略

带状态的路径图

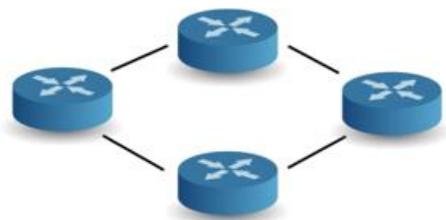
OSPF综合实现样例

对OSPF协议建模，每条物理链路抽象为两条有向边，有向边的权值表示cost属性。

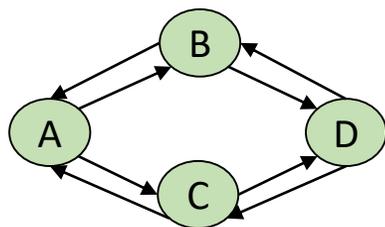
P_{12one} 表示节点1到节点2的最优路径， $P_{12others}$ 表示其它路径， $Cost(p)$ 表示路径p链路cost的累加值。

OSPF指定最优路径问题转化为有向边权值为变量的不等式组求解出一组可行解的问题。

$$Cost(P_{12one}) < Cost(P_{12others})$$



建模



从A到B的流量走A,C,D,B
从B到C的流量走B,D,C

转发需求

$$AC+CD+DB < AB$$
$$BD+DC < BA+AC$$

不等式组

$$AC=1 \quad CD=1 \quad DB=1 \quad AB=4$$
$$BA=2 \quad BD=1 \quad DC=1 \quad CA=1$$

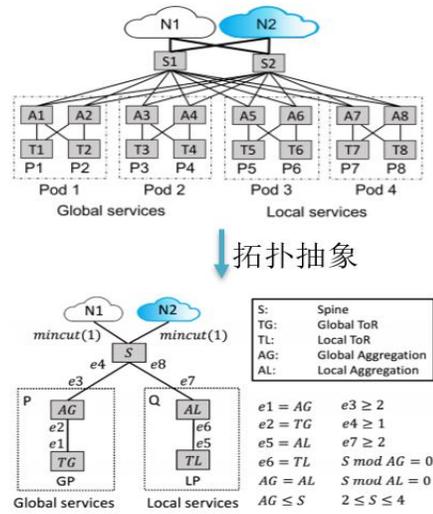
一组可行解

求解器

网络配置综合研究-1

- 绿地网络：使用DSL/形式化语言完备描述用户配置需求，通过推理计算求解出每个设备配置的关键参数。
- 明确需求范围（骨干网/数据中心网络，网络协议，安全策略等），根据需求范围设计DSL及综合方法。

使用正则表达式和确定有限自动机构建的PG图实现相关参数计算，自动生成BGP配置。



(a) Network Specification N

```

Pvd(TC, Router, NextHop) :-
  Route(TC, Router, NextHop, Proto),
  SetAD(Proto, Router, Cost)
minAD(TC, Router, Cost)
minAD(TC, Router, minCost) :-
  Route(TC, Router, NextHop, Proto),
  SetAD(Proto, Router, Cost)
Route(TC, Router, Next, "static") :-
  SetStatic(TC, Router, NextHop)
Route(TC, Router, NextHop, "ospf") :-
  BestOSPFRoute(TC, Router, NextHop)
    
```

(b) Topology

(c) Requirements φ_R

```

Path(N1, A, [A,B,C,D])
Path(N2, A, [A,D])
Path(Ext, A, [A,C])
Path(Ext, D, [D,B])
    
```

(e) Configuration for Router A

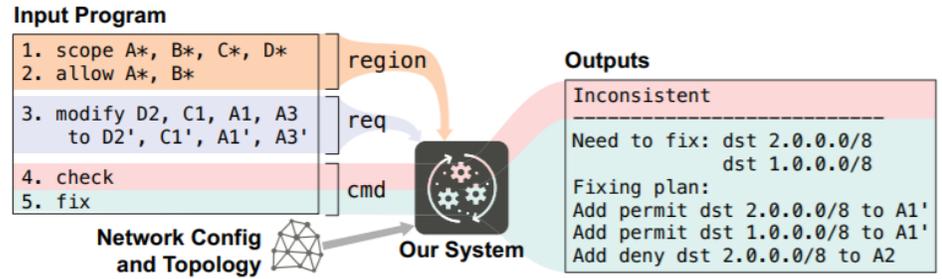
```

! 10G interface to B
interface TenGigabitEthernet1/1/1
...
ip ospf cost 10
! 10G interface to C
interface TenGigabitEthernet1/1/2
...
ip ospf cost 5
...
! static route to B
ip route 10.0.0.0 255.255.255.0 130.0.1.2
    
```

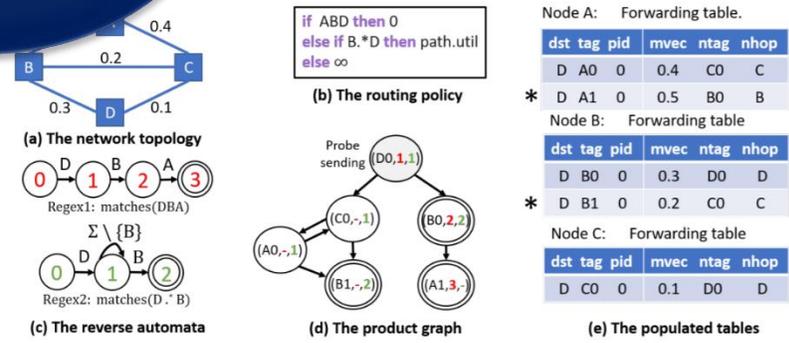
将配置求解问题转化为求解Datalog问题的输入问题，在层次迭代中使用SMT求解，自动生成BGP、OSPF和静态路由配置。



设计高层次的ACL更新策略描述语言，通过建模对ACL更新的安全性进行分析生成配置规划。

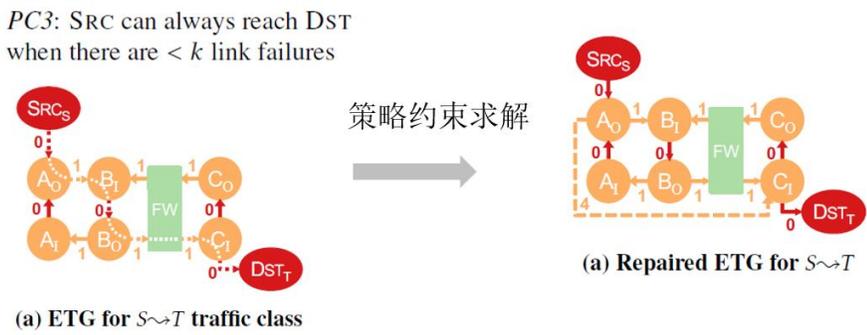


使用探针对网络链路状态进行收集，使用PG图进行路径提取及参数赋值，自动生成P4代码。



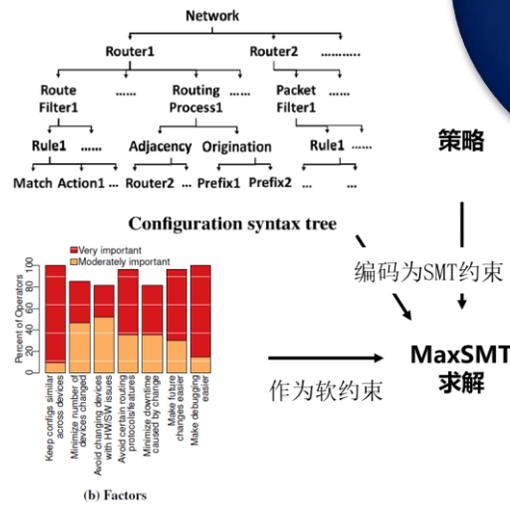
网络配置综合研究-2

- 棕地网络：使用形式化方法对配置建模并编码，通过SMT等求解出满足用户策略的配置更新的关键参数。
- 明确需求范围（拓扑，管理目标，用户策略等），根据需求采用对应建模和优化方法进行求解。

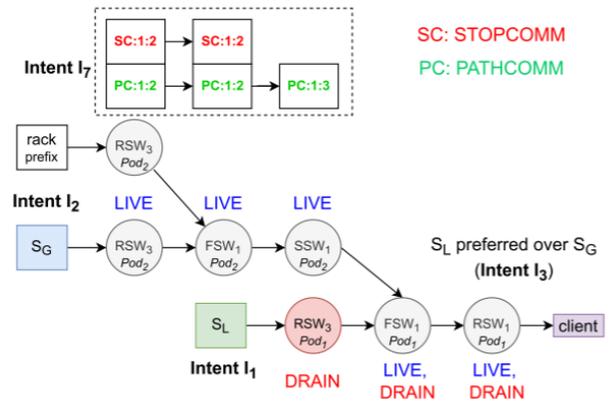


使用控制平面语义编码和基于MaxSMT的约束求解，可以自动生成正确的、最小的控制平面修复。

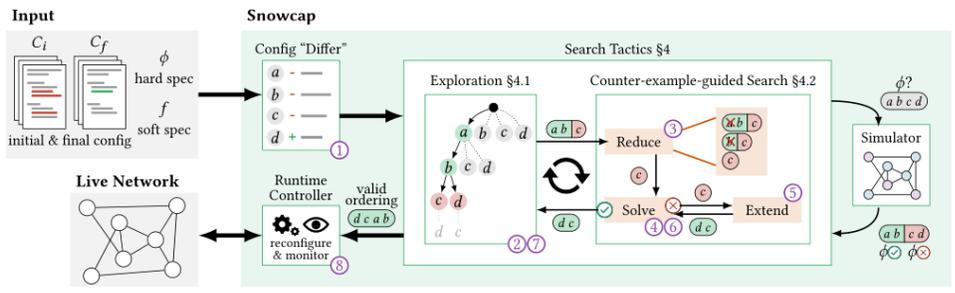
将配置和潜在更新编码为类似语法树的基于MaxSMT的模型，并将管理目标编码为软约束。最终得到涵盖多个管理目标的更新。



设计一个用于表达网络行为的高级语言及自动生成配置的编译器，以支持频繁的策略和网络变更。可以生成并部署多个并行策略集合。



将配置更新问题表述为约束条件下的优化问题，提出了一种满足软硬约束的综合框架，并采用一种反例引导搜索来优化查找修复并进行配置部署。



CPR
SOSP 2017

Aura
NSDI 2023

AED
CoNEXT 2020

Snowcap
SIGCOMM 2021



Part 02

配置综合实践

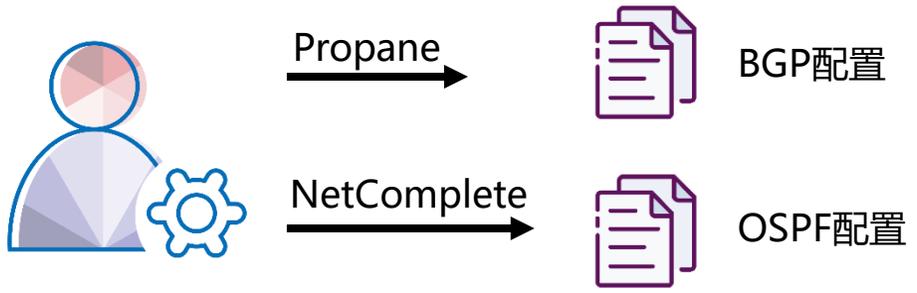


跨层多协议配置综合-Lynx

研究动机

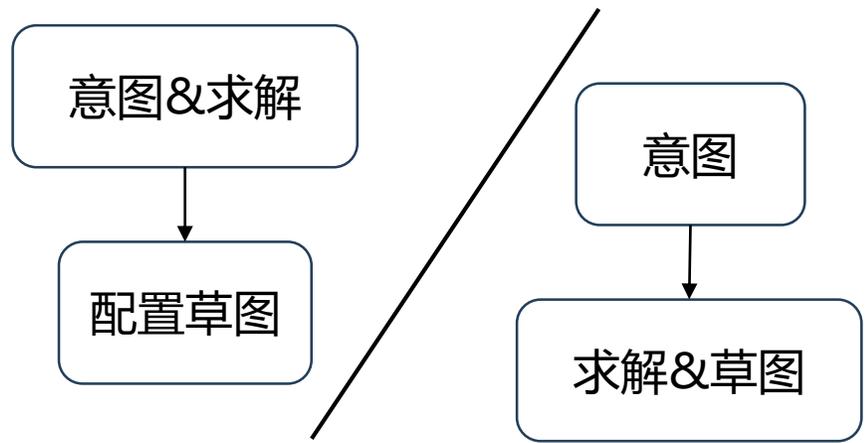
Lynx: Cross-Layer Multi-Protocol Network Configuration Synthesis.

➤ 现有配置综合工具不支持复杂网络场景



- 支持协议单一
- 支持策略以路径控制策略为主
- 缺少端到端的自动配置能力

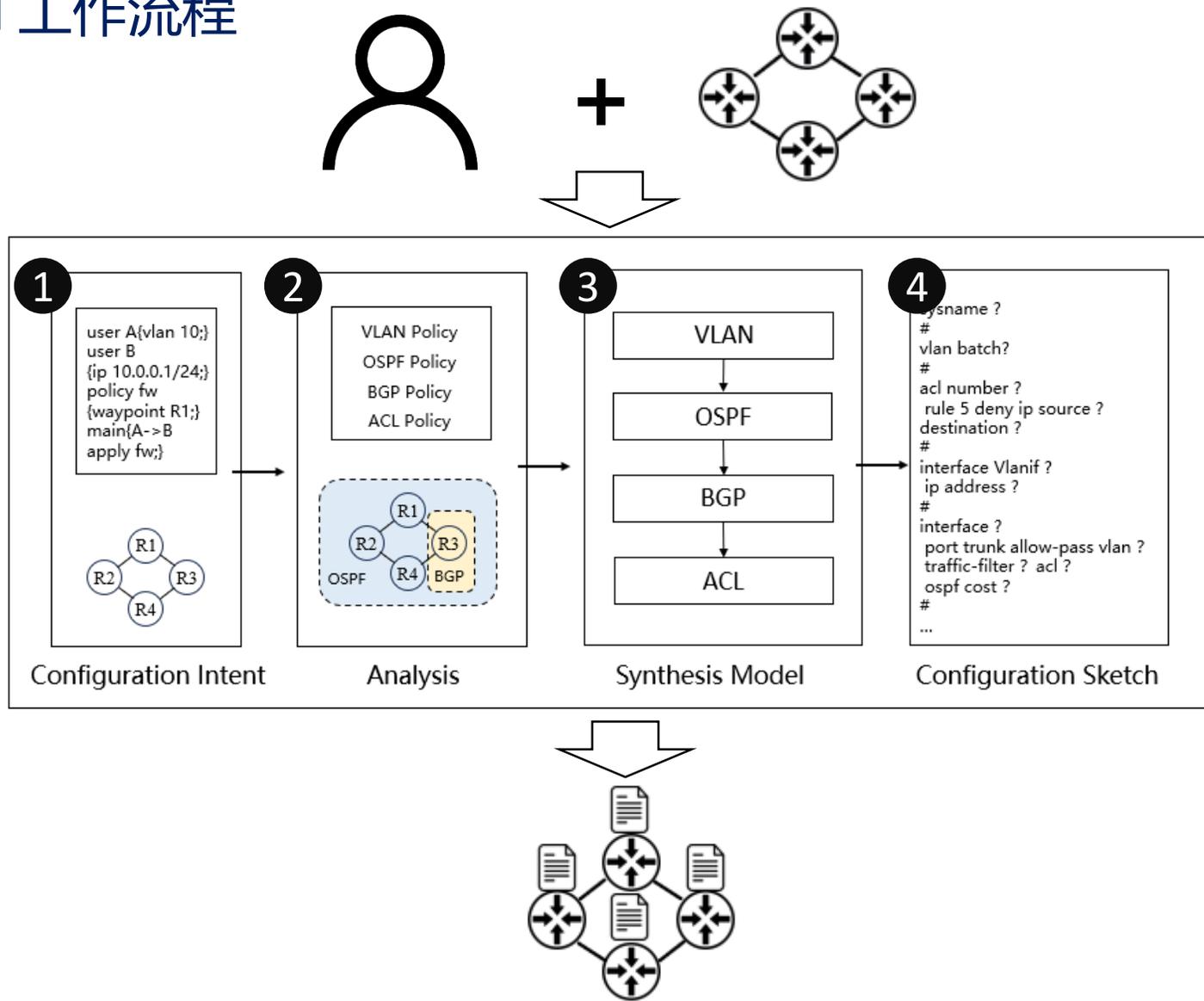
➤ 现有配置综合方法拓展困难



- 协议求解方法与其他模块高度耦合

跨层多协议配置综合-Lynx

□ 工作流程



- 1** 将策略与用户解耦，策略以系统关键字语句结构进行描述，方便拓展。
- 2** 根据意图与拓扑信息，划分各协议配置范围并将高层次意图分解为协议级别的策略。
- 3** 按协议依赖关系依次对VLAN、OSPF、BGP和ACL配置参数进行求解。
- 4** 将配置参数映射到对应设备型号的配置草图中。

跨层多协议配置综合-Lynx

□ 性能分析

TABLE I

NUMBER OF TOPOLOGY SIZES, POLICY TYPE AND SOLUTION TIMES FOR LYNX SYNTHESIS

Network size	Policy type	4 policies	8 policies	16 policies	32 policies	64 policies
Small (7-31)	TE	3.67	7.46	14.83	29.53	58.10
	Isolation	0.92	1.07	1.35	1.87	2.34
	Mix	4.39	6.21	10.77	20.35	47.09
Medium (35-63)	TE	8.03	17.38	39.30	87.86	151.77
	Isolation	15.46	19.42	28.13	53.31	121.80
	Mix	13.83	21.36	35.13	71.42	154.32
Large (63-94)	TE	14.03	29.02	72.58	173.17	396.97
	Isolation	40.12	68.32	124.91	257.35	554.01
	Mix	18.56	63.50	102.60	266.39	428.46

1

路径控制相关策略，为多种策略与多种协议求解混合后的时间。小型网络时间小于1min，大型网络时间小于10min。

2

ACL隔离策略，随机混合多种偏好，因其全部使用形式化方法，其时间开销增长大于路径控制策略。小型网络时间小于10s，大型网络时间小于10min。

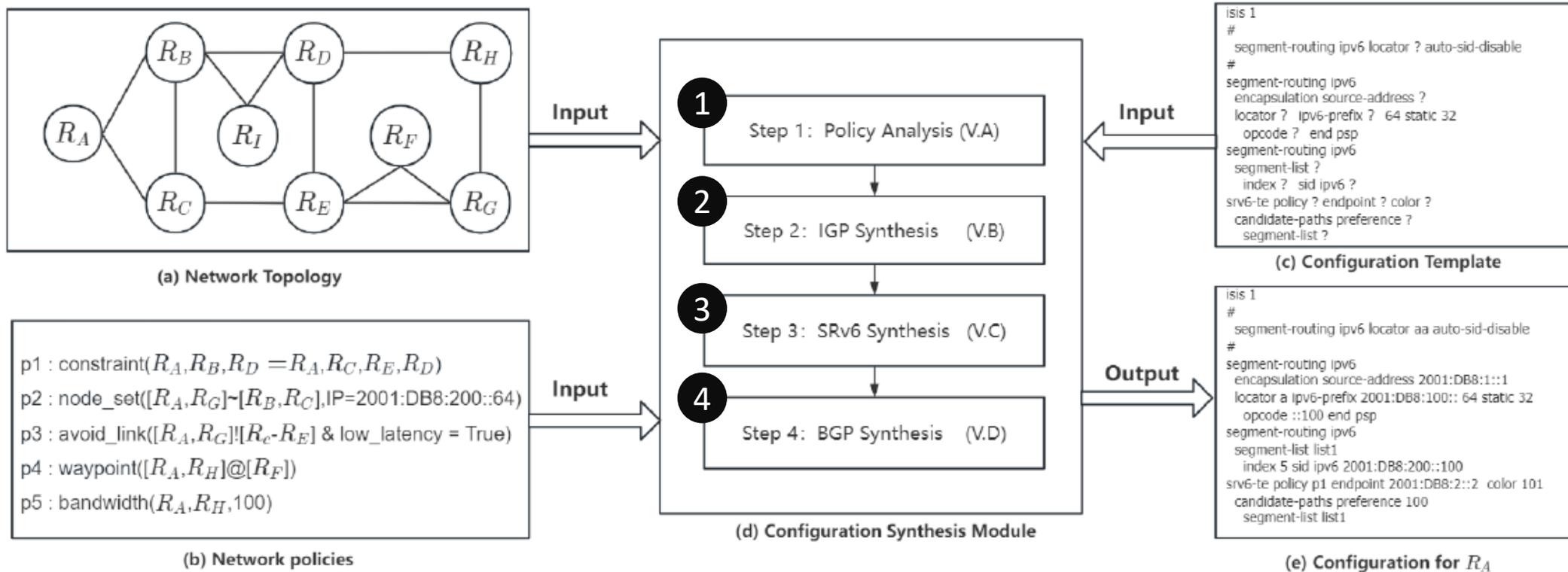
3

随机路径控制与隔离策略，其时间开销整体介于路径控制与隔离策略间。

SRv6配置综合-AutoSRv6

□ 工作流程

AutoSRv6: Configuration Synthesis for Segment Routing over IPv6.



- 1 根据策略类型使用不同的协议模型进行综合。
- 3 SRv6实现航路点，显式路径和带宽约束等策略。

- 2 基于IGP实现负载均衡策略。
- 4 基于BGP实现自动引流。

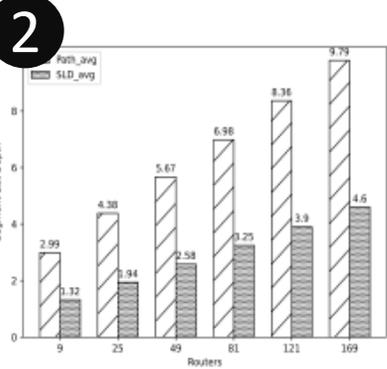
SRv6配置综合-AutoSRv6

性能分析

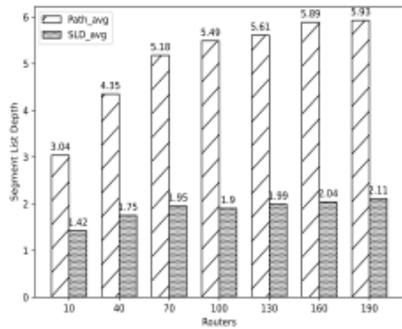
1 TABLE IV: Number of topology sizes, requirement type and solution times for SRv6 Synthesis

Network size	Requirement type	4 requirements	8 requirements	16 requirements	32 requirements	64 requirements
Small (9-25)	Waypoints	0.34	1.84	4.62	12.42	24.41
	Isolation	0.26	1.07	3.12	9.23	13.54
	Bandwidth	0.60	3.82	5.03	10.86	21.98
	TE	0.21	0.70	0.87	1.09	2.51
	Mix	0.79	4.21	4.73	14.53	27.09
Medium (36-64)	Waypoints	6.23	9.62	27.30	81.86	132.78
	Isolation	0.47	5.15	8.33	23.47	41.24
	Bandwidth	4.36	6.17	16.99	64.88	93.03
	TE	2.36	7.13	14.29	31.46	61.3
	Mix	27.83	47.36	65.13	91.42	136.32
Large (65-81)	Waypoints	14.73	26.02	58.28	104.43	209.71
	Isolation	4.87	17.92	34.91	56.44	146.54
	Bandwidth	6.39	27.50	46.68	91.39	201.46
	TE	3.16	9.98	26.22	51.92	128.58
	Mix	39.61	63.24	108.97	194.46	335.48

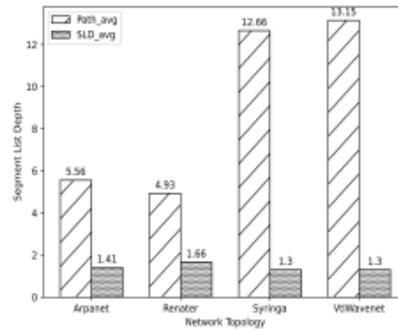
1 不同网络规模，不同策略数量下各流量策略的时间开销。小型网络时间小于1min，大型网络时间小于10min。



(a) Regular grid network topology.



(b) Networkx random network topology.



(c) Real network topology.

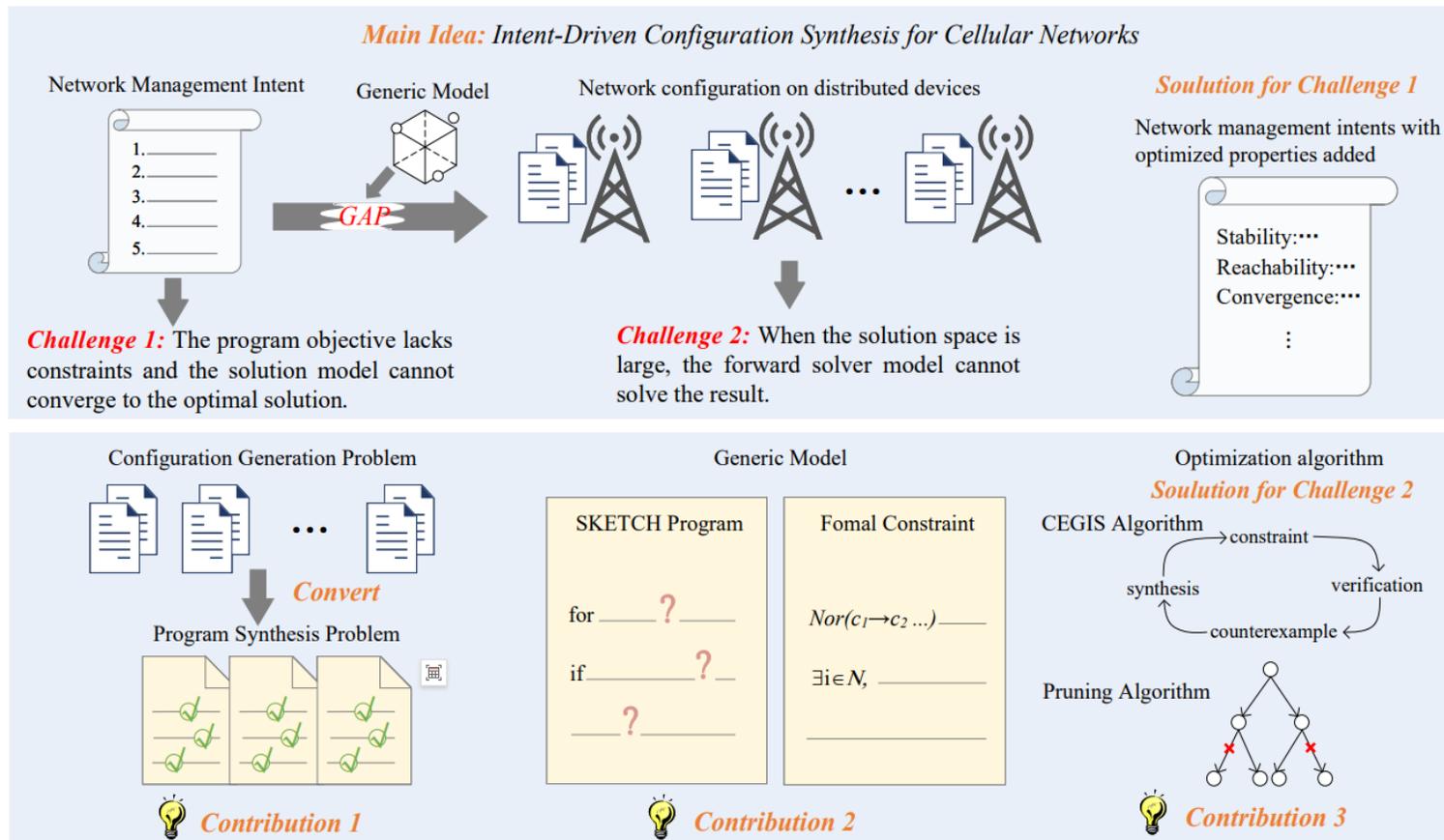
2 不同网络下的Segment list长度优化效果，长度缩短约2x-10x。

蜂窝网络配置综合-Drone

◆ Drone: Intent-Driven Cellular Configuration Generation Using Program Synthesis, *IEEE JSAC*, 2023.

主要贡献

- 实现一个从意图到配置的智能化蜂窝网络配置综合框架;
- 采用形式化综合方法求解出完备且精确的蜂窝网络配置参数;
- 构建通用的蜂窝网络配置语义模型, 弥补意图与配置之间的语义鸿沟。



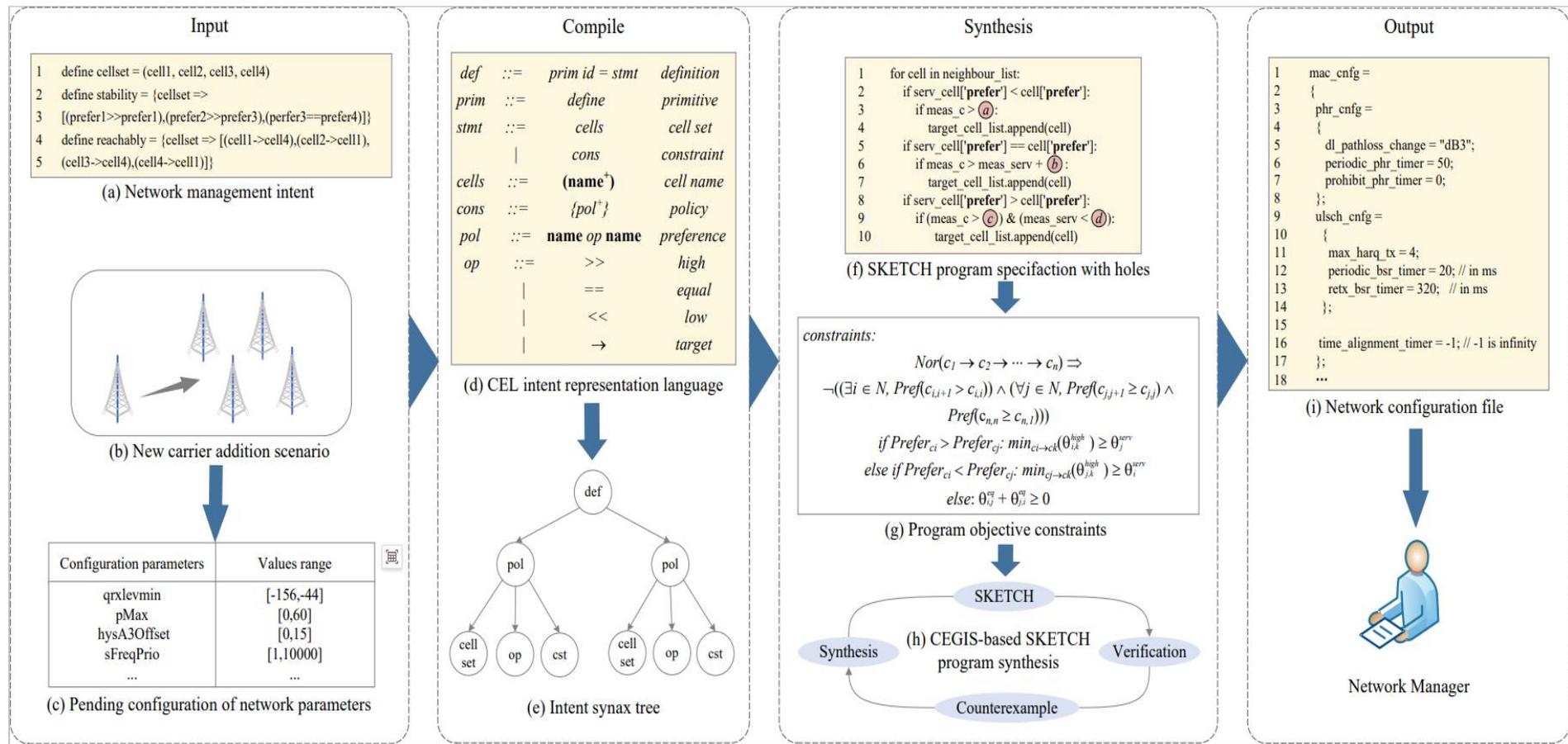
蜂窝网络配置综合-Drone

□ **输入**：设计一种**高级意图语言**，用于表达运营商的对于蜂窝网络的高级管理意图。

□ **编译**：将输入的意图语言编译成可以灵活操作的**中间表示**。

□ **综合**：构建通用的蜂窝网络功能的**配置程序模型及程序目标的形式化约束**，应用程序综合方法进行求解。

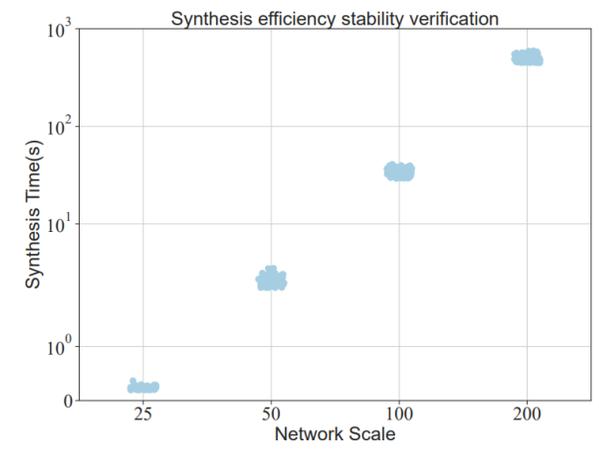
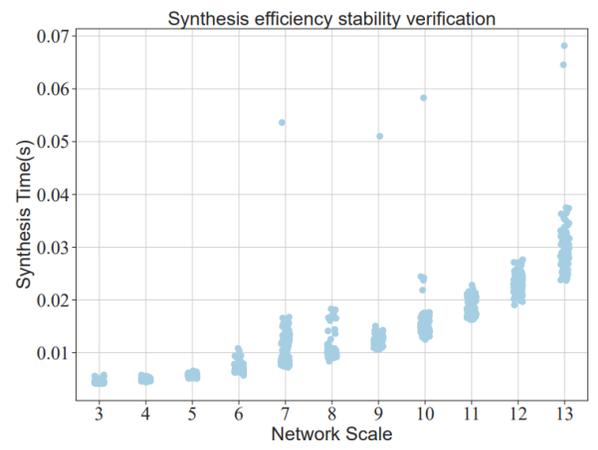
□ **输出**：最终的蜂窝网络配置文件。



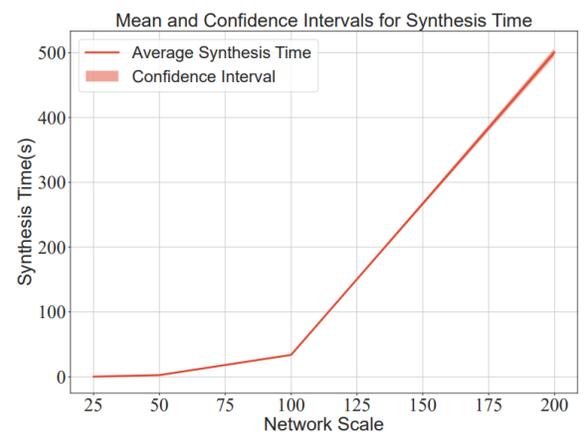
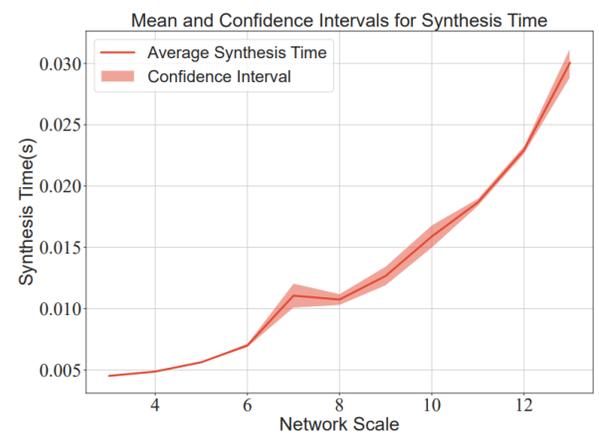
蜂窝网络配置综合-Drone

蜂窝网络配置综合结果

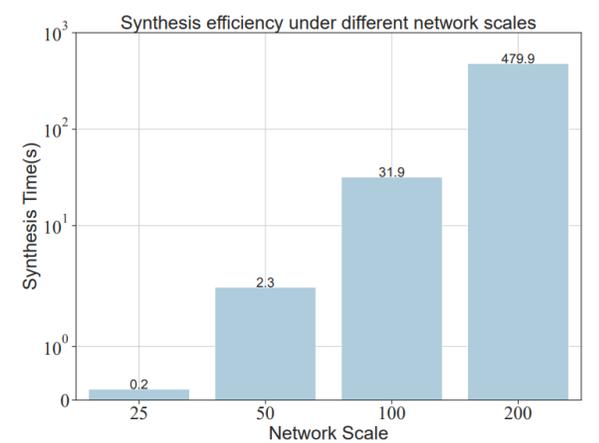
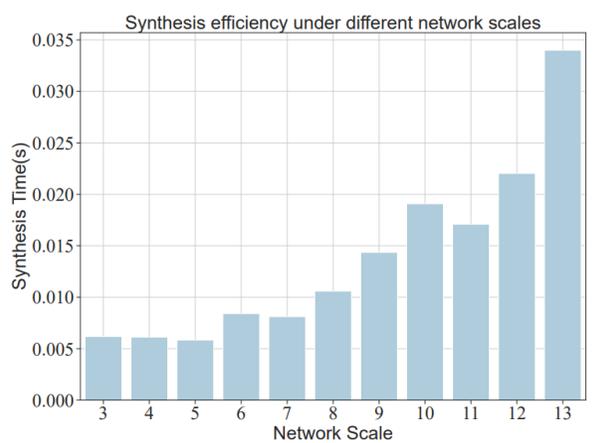
综合系统在小规模网络拓扑下综合时间小于1s，大规模拓扑下不超过10min。优化算法将综合速度提升3.8-485倍。



(a) Synthesis Efficiency Stability Verification



(b) Mean and Confidence Intervals for Synthesis Time



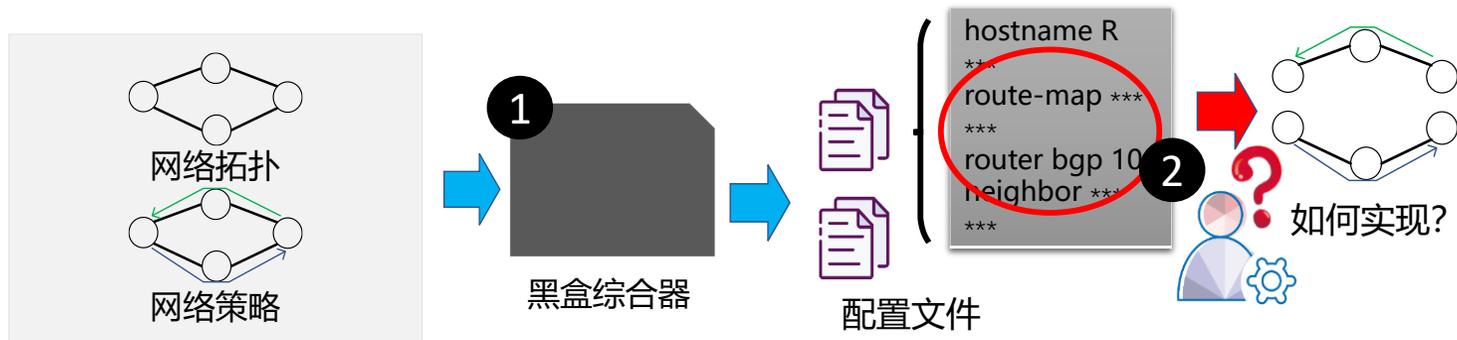
(c) Synthesis Efficiency under Different Network Scales

可解释的配置综合技术-NetCR

□ 现有配置综合工具缺乏可解释性

NetCR: Knowledge Graph based Recommendation Framework for Manual Network Configuration. *IEEE IoT Journal*.

➤ 综合过程缺乏可解释性



① 分析过程是不透明，呈现**黑盒特征**，导致不知道分析是否可靠和全面。

□ 例如：配置综合工具CPR在K链路故障分析下使用了错误的ACL模型。

② 黑盒掩盖了实施的高级策略和配置之间的一致性关系。

➤ 综合结果缺乏可解释性

• 综合的配置不被信任

- 生成的配置内容未知。
- 需要管理员自行阅读全网配置。
- 依赖专家经验且人工易出错。

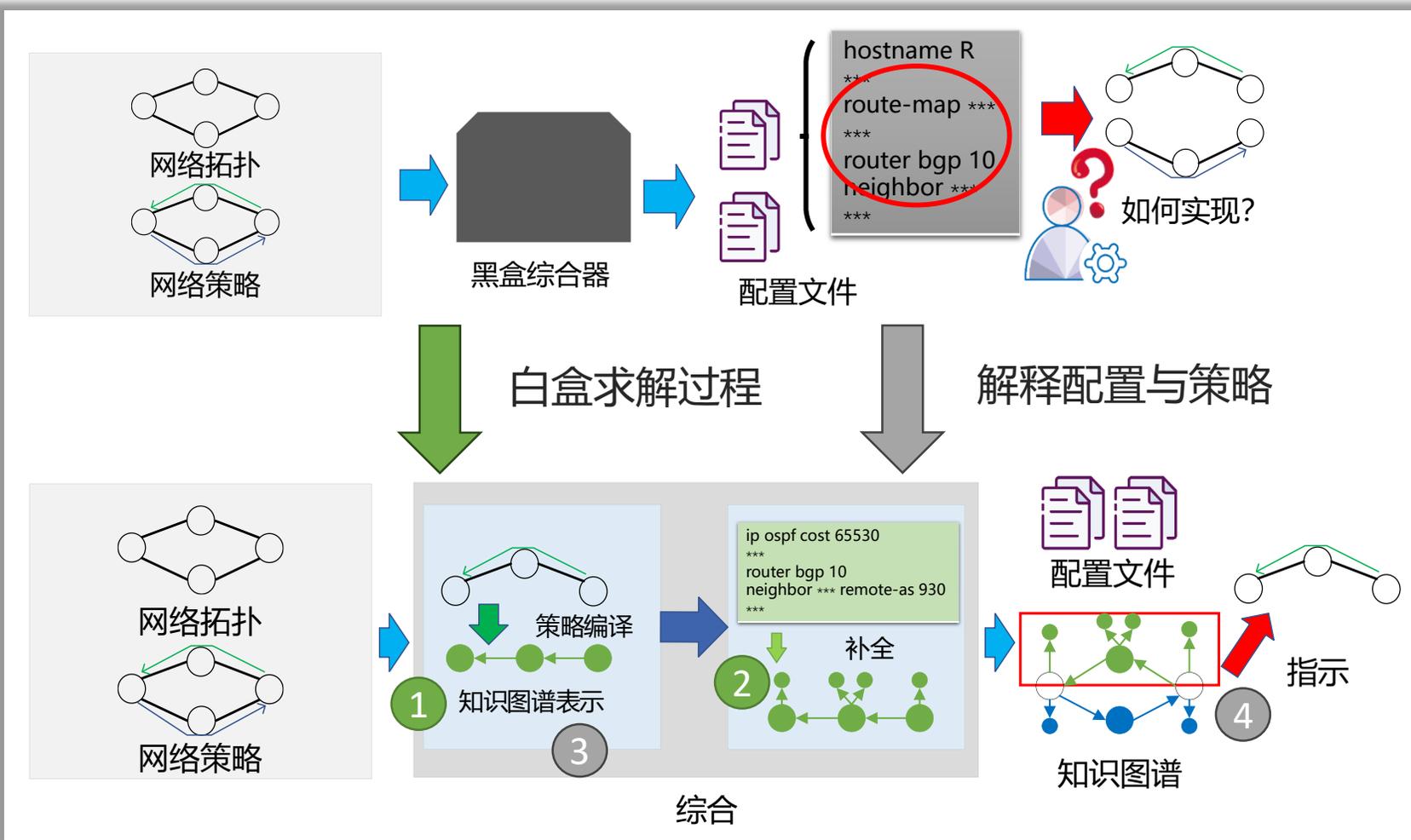


• 实际配置仍以人工模板和脚本为主

- 人工编写模板。
- 掌握所有配置内容。
- 脚本批量生成。

可解释的配置综合技术-NetCR

□ 实现方案：基于知识图谱的配置综合工具NetCR



➤ 基于知识图谱实现白盒综合

- 1 知识图谱作为中间表示建模配置。
- 2 将综合问题转换为配置推理与补全问题。

➤ 基于知识图谱提供可解释性

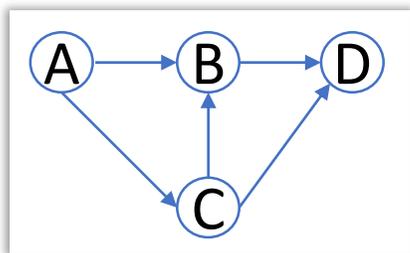
- 3 解释策略的求解过程。
- 4 解释高级策略和配置之间的一致性关系。

可解释的配置综合技术-NetCR

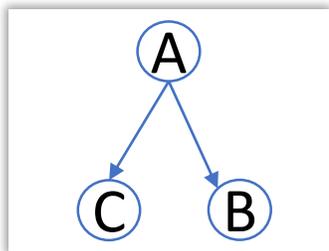
知识图谱的可解释性

- 知识图谱的每个**顶点**和**边**都能作为解释内容，**推理分析**过程即**解释**过程。
- 灵活的可解释范围。

- 通过生成任意类型子图解释内容

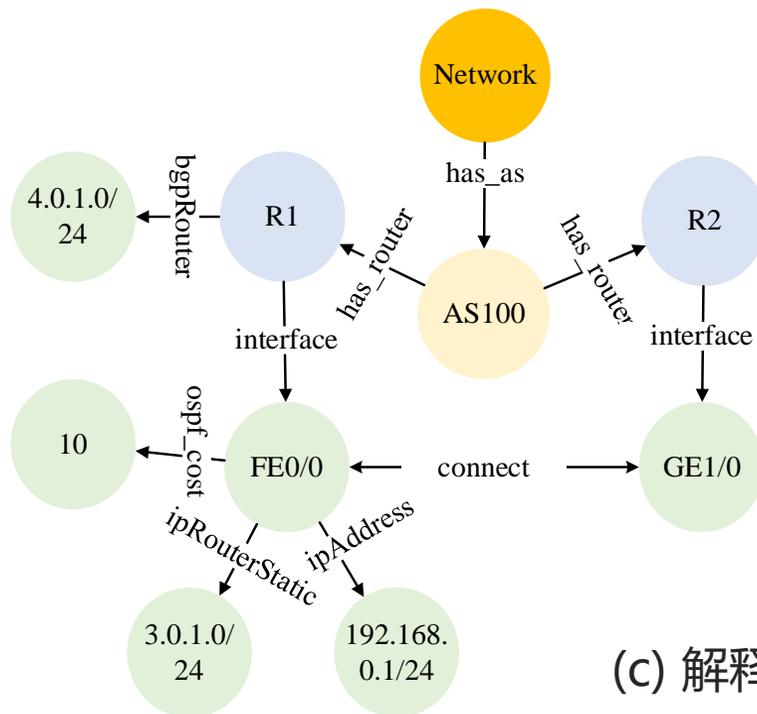


(a) 抽象图



(b) 树

- 通过标签生成详细的解释内容



(c) 解释R1与R2

可解释的配置推荐技术-ConfigReco

研究动机

ConfigReco: Network Configuration Recommendation with Graph Neural Networks. *IEEE Network Magazine*.

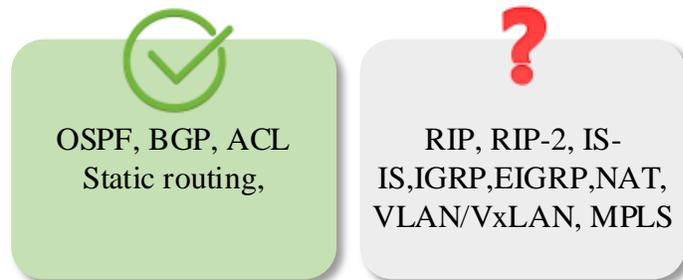
现有综合工具使用受限

a. 缺乏可解释性

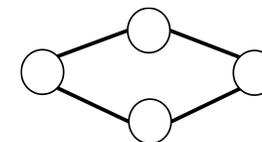


黑盒综合器

b. 通用性有限



b. 伸缩性有限

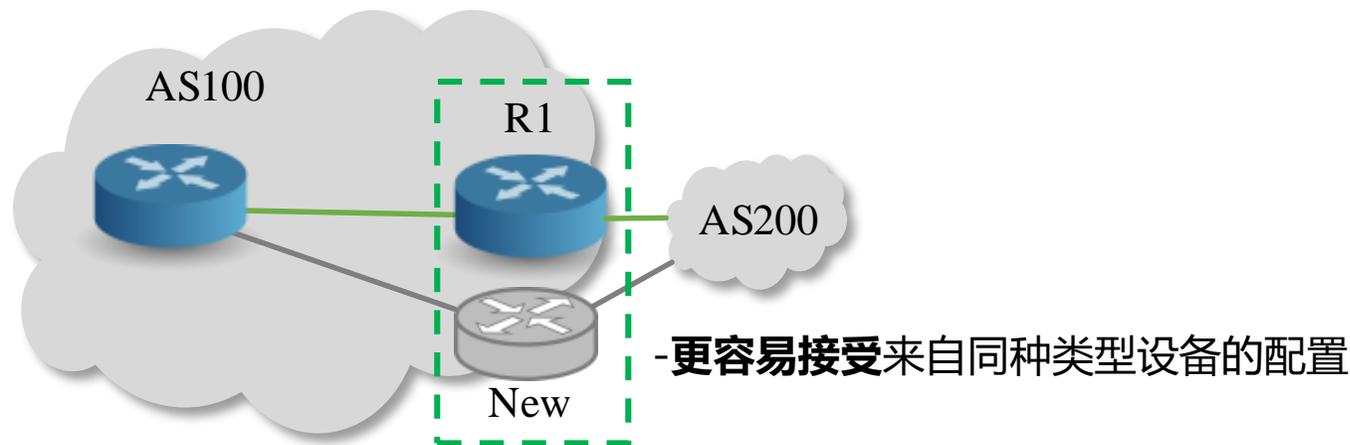


64 nodes

- SyNET > 24小时
- NetComplete > 6小时

配置相似性-配置推荐的基础

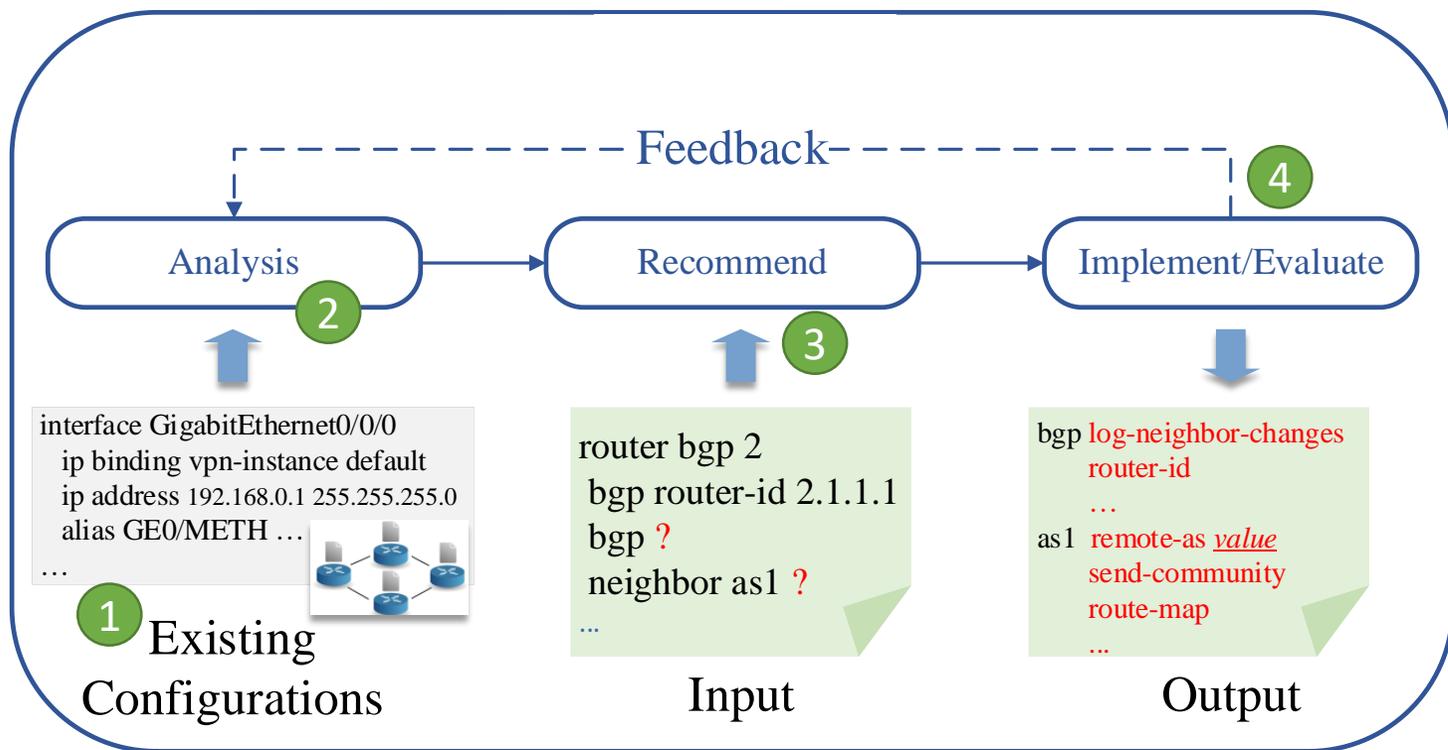
- 人工编写统一配置模板
- “复制-粘贴”是配置的常用手段



可解释的配置推荐技术-ConfigReco

□ 实现方案

➤ 基于知识图谱的配置推荐工具ConfigReco

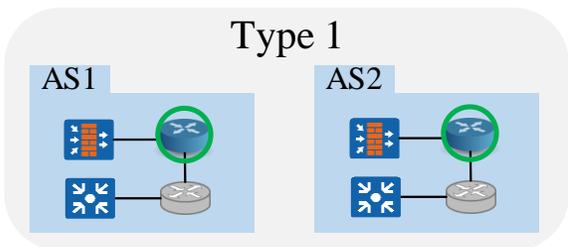


- 1 知识图谱将配置文件转换为图结构。
- 2 图算法分析配置命令之间的依赖关系。
- 3 根据输入配置上下文与依赖关系推荐配置。
- 4 实时反馈机制保证推荐精度。

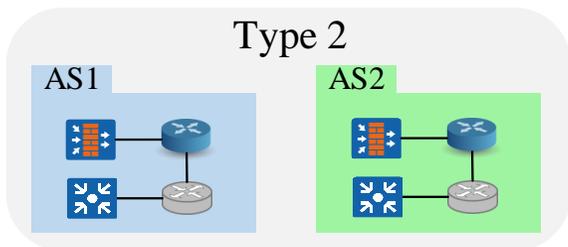
可解释的配置推荐技术-ConfigReco

➤ 支持不同范围的“配置相似性”

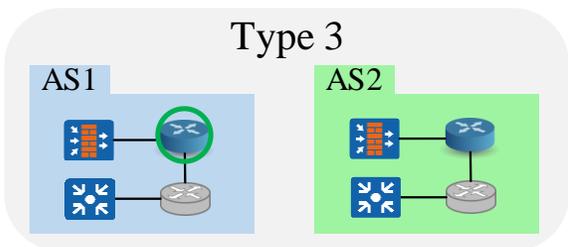
○ 推荐范围



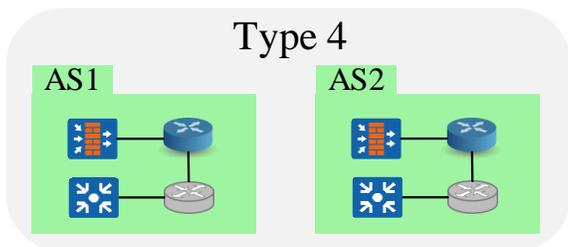
(a) 设备推荐



(b) 自治域推荐



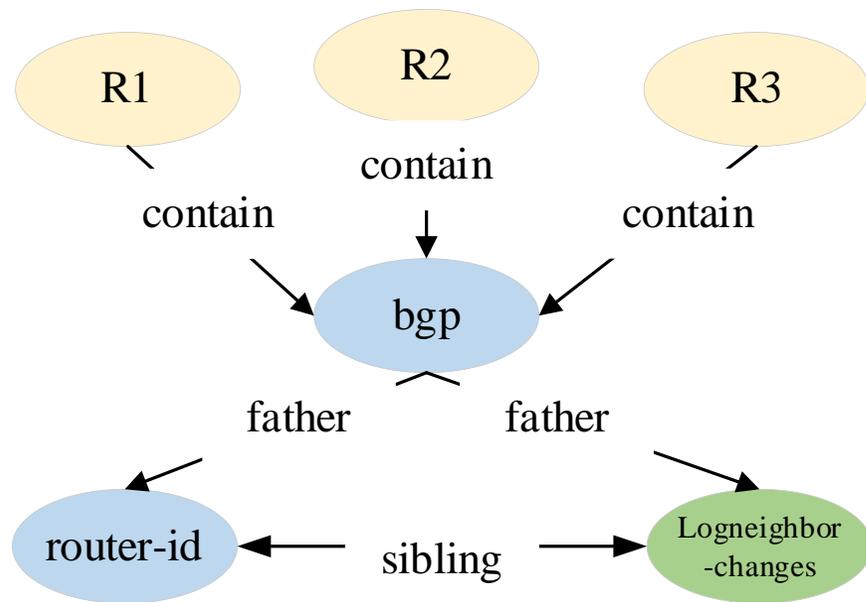
(c) 混合推荐



(d) 全局推荐

➤ 知识图谱解释推荐原因

- 配置的所在位置，如路由器
- 配置的上下文环境





Part 03

配置翻译



基于机器翻译的配置翻译

背景与动机

- 真实场景下的多云多厂商网络存在异构性
- 异构网络的配置迁移和备份工作是繁琐困难的

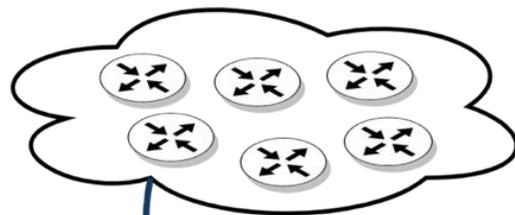
网络配置翻译是为了解决网络配置管理过程中的配置迁移等问题而提出的。然而配置翻译技术存在着一些问题：

- 基于规则/模板的配置翻译可扩展性差
- 网络配置领域无高质量标注监督数据

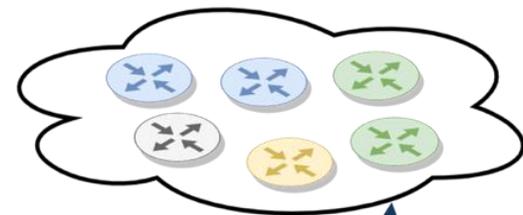
在自然语言和编程语言翻译领域，神经机器翻译技术（NMT）取得了较大进展。将NMT应用到网络配置迁移过程中可以大大加速配置迁移过程。

Opportunities and Implementation of Neural Machine Translation in Network Configuration. *IEEE Network Magazine*.

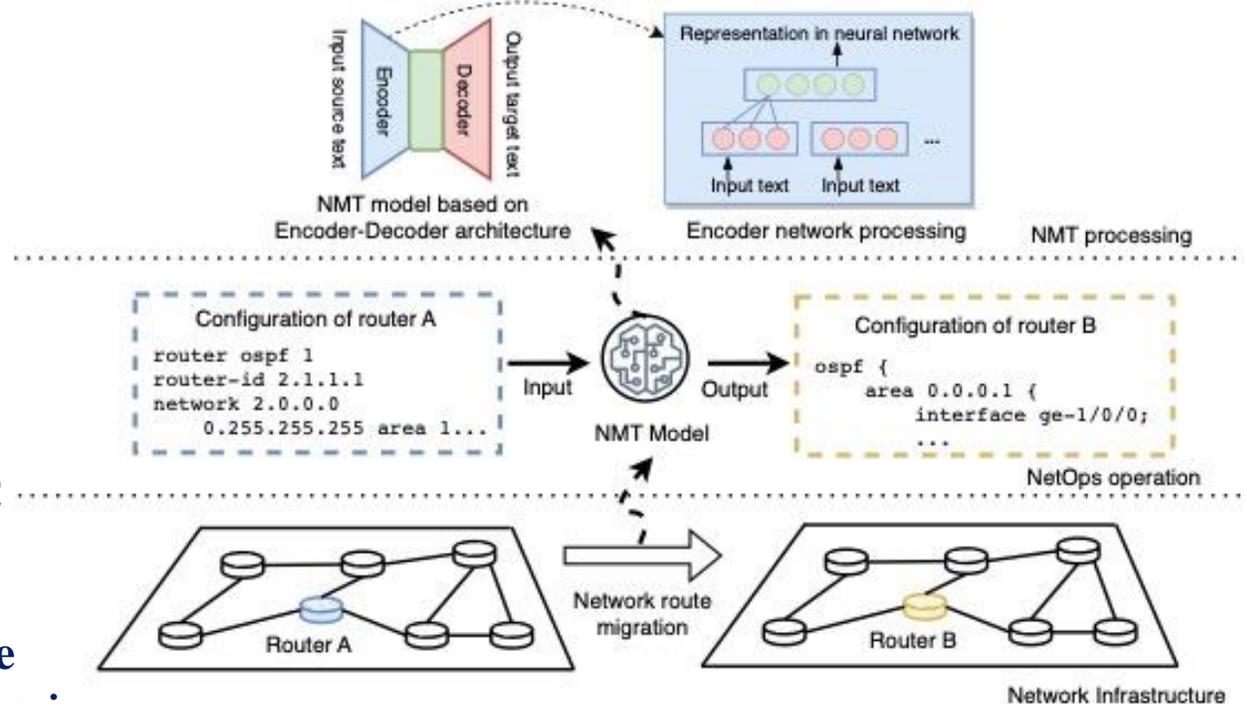
理想下的设备均一网络



真实场景下的异构网络



多供应商之间的竞争和新旧设备共存



基于机器翻译的配置翻译

无监督网络配置翻译方法

无监督配置训练方法：针对网络配置领域无标注数据特点，采用无监督方法实现了基于机器翻译的配置翻译方法。我们使用一个seq2seq模型，该模型由一个基于Transformer架构的编码器译码器组成，对所有厂商配置语言都使用相同的这一个的共享模型^{[Cisco][Juniper]}。使用无监督机器翻译的三个原则来训练它，即**初始化、语言建模和反向翻译**。

(1) 跨厂商配置掩码配置模型预训练



表示相同指令的配置片段会映射到相同的表示
(建立了厂商无关模型)

(2) 噪声自编码



训练解码器始终生成有效的序列，提高鲁棒性（中间表示到目标组件的生成）

(3) 反向翻译



使得模型生成可用于训练的并行数据，形成自监督模型，提升翻译效果

基于机器翻译的配置翻译

无监督网络配置翻译方法实验和表现

我们构建了一个超过1G的非并行原始配置数据集，并对其应用中所示的网络配置预处理框架，进行数据增强等操作以满足模型输入需求。

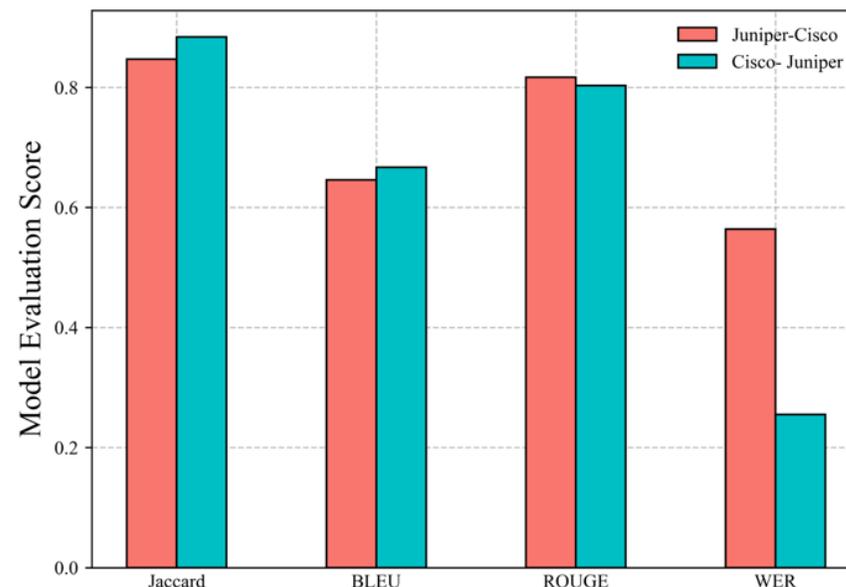
同时我们构建了100对测试配置用于我们模型的评估，这些数据来源于真实的厂商用户手册和网络社区。评估效果如图所示，给出了一个翻译案例并在多个指标上测试了模型效果。

Module	Main function
config_extract	Parse the user manual and extract the network configuration
config_clean	Delete non-configuration text information
config_enhancement	Randomly change, delete, or add configuration information to enhance the configuration text
config_tokenize	Split the network configuration text into individual words or symbols
config_quality	The number of various configuration parameters is collected to visualize the overall data set

TABLE 1. Network configuration preprocessing framework modules and functions.

Input	Output
<pre>bgp { group GROUP1 { type external; local-address 2.1.1.1; family inet { unicast; } import POL; local-as 2; neighbor 10.12.11.1 { description R1; peer-as 1; } } }</pre>	<pre>router bgp 10 bgp router-id 2.1.1.1 neighbor GROUP1 peer-group neighbor GROUP1 remote-as 10 neighbor 10.12.11.1 remote-as 30 neighbor 10.12.11.1 description R1 ! address-family ipv4 bgp dampening bgp bgp additional-paths select all bgp additional-paths send receive exit-address-family</pre>

(a) Translation Examples



(b) Evaluation Indicators



Part 04

配置大模型



网络配置生成、翻译和摘要

ConfigAid: A Pre-trained Model for Configuration Generation Tasks.

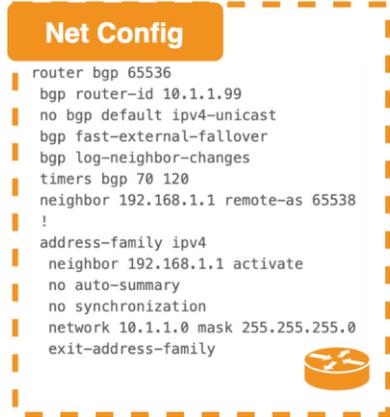
背景



It's easy. I've learned it before.



LM



I'm sorry, I don't understand...

LLM不能准确理解配置知识

Test 1: 通用语言模型对配置语句的支持和理解能力

z To create a set clause to apply the cost community attribute to routes that pass through a route map, use the `xxx` command in route-map configuration mode.

To create a set clause to apply the cost community attribute to routes that pass through a route map, use the `set community` command in route-map configuration mode.

set extcommunity cost ✓

Dataset	Supported statement	Unsupported statement
326	119	207

表1: GPT配置语句支持能力测试结果

Test 2: 通用语言模型在配置生成方面的能力

Dataset	Num
Small	150
Medium	150
Large	150

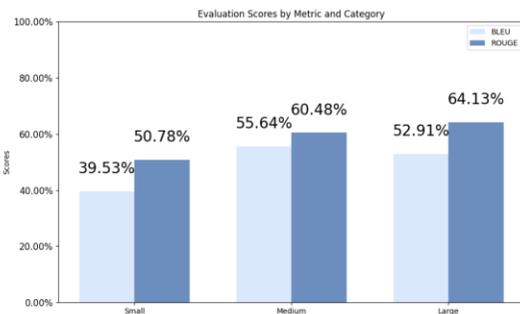
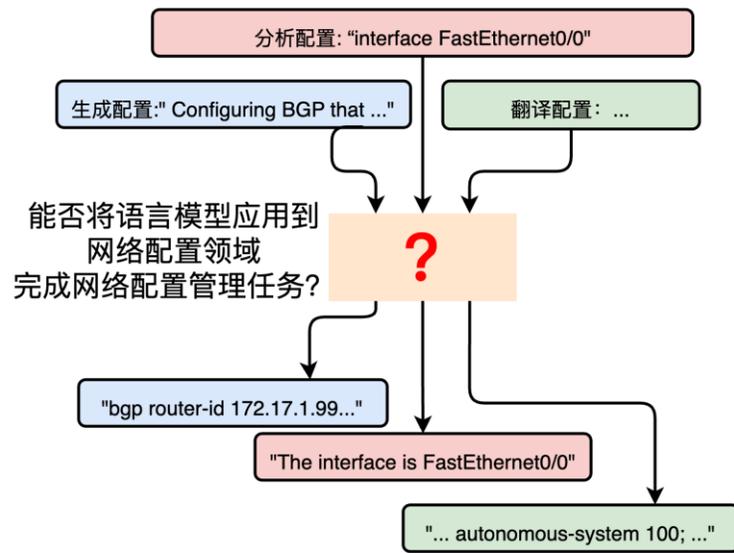
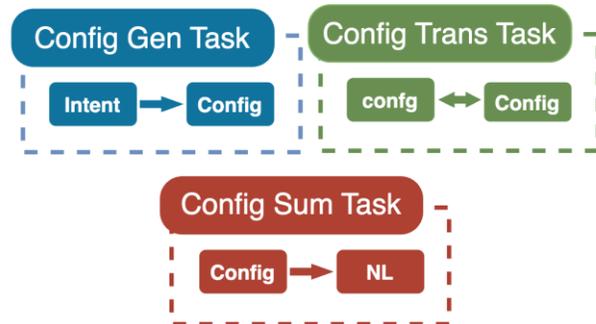


图2: GPT配置生成能力测试结果

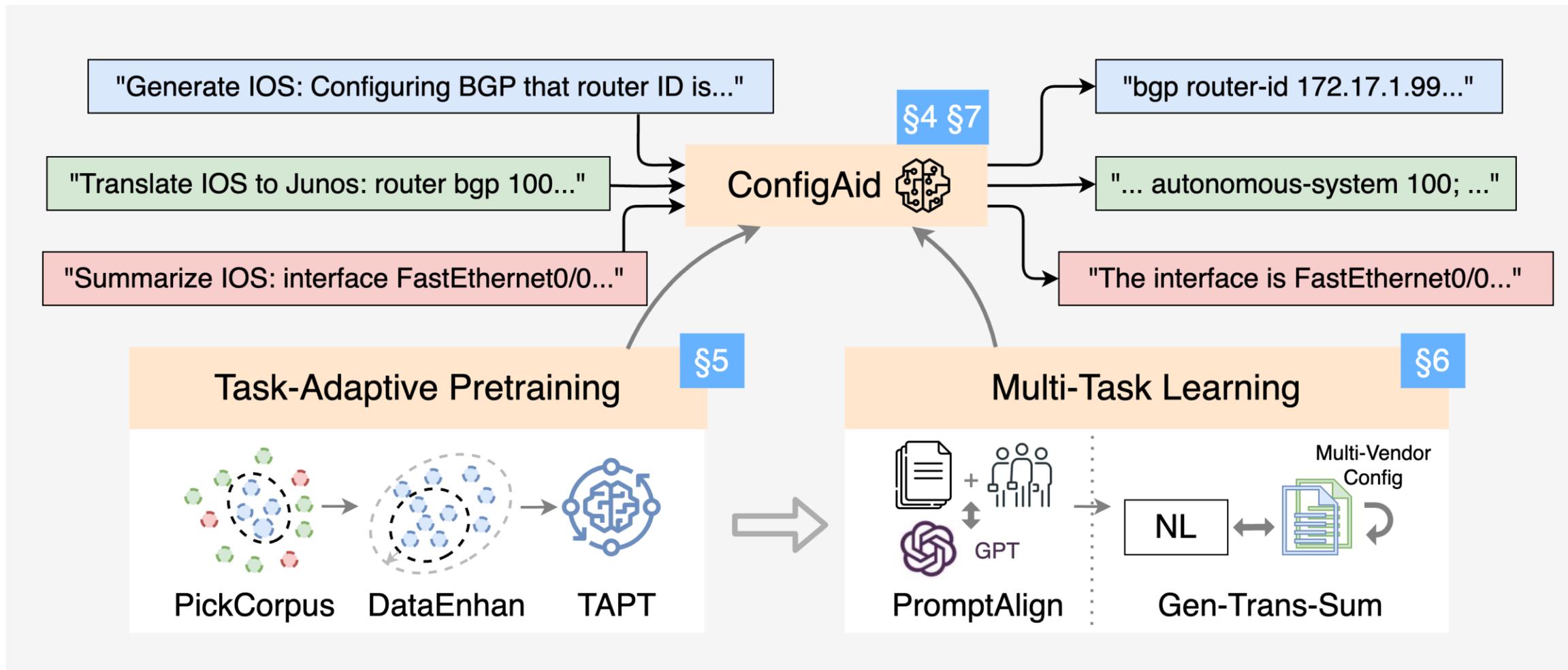
GPT在网络配置领域表现不佳



如何使用LLM完成配置任务

网络配置生成、翻译和摘要

□ ConfigAid : 配置生成、翻译和摘要工具



网络配置生成、翻译和摘要

ConfigAid 效果评估

```
1 ip community-list standard comm1 permit 1:2 1:3
2 ip prefix-list pfx permit 192.168.2.0/24
3 route-map RMO permit 10
4 match community comm1
5 match ip address prefix-list pfx
6 set local-preference 200
7 route-map RMO permit 20
8 set metric 90
```

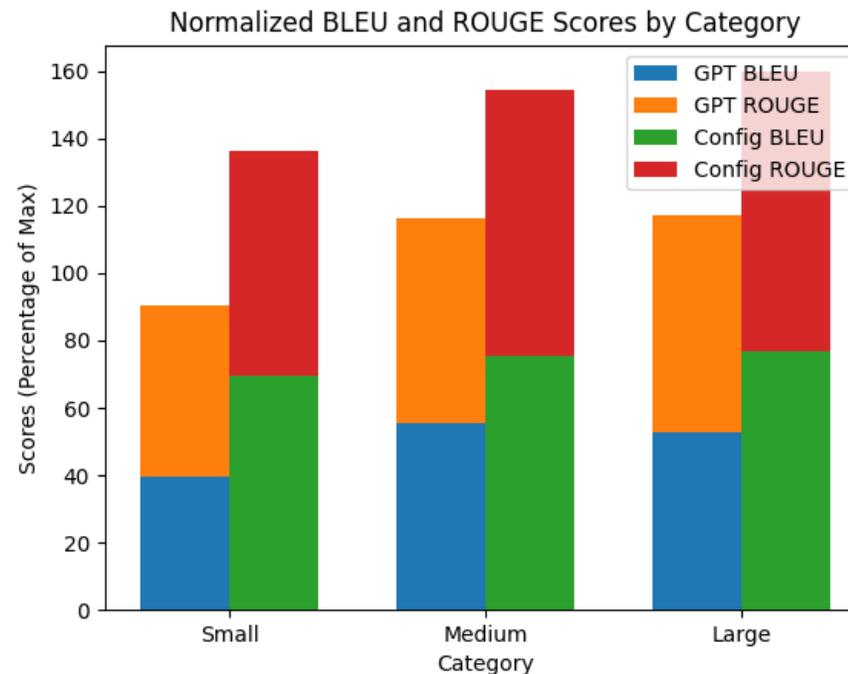
真值

```
1 ip community-list standard comm1 permit 1:2 permit 1:3
2 ip prefix-list pfx permit 192.168.2.0 255.255.255.0
3 route-map RMO permit 10
4 match community comm1
5 match prefix-list pfx
6 set local-preference 200
7 route-map RMO permit 20
8 set metric 90
```

GPT3.5生成效果

```
1 ip community-list comm1 permit 1:2 1:3
2 ip prefix-list pfx permit 192.168.2.0/24
3 route-map RMO permit 10
4 match community comm1
5 match ip address prefix-list pfx
6 set local-preference 200
7 route-map RMO permit 20
8 set metric 90
```

ConfigAid生成效果



ConfigAid的配置生成表现优于GPT3.5

指标	BLUE	ROUGE-1	ROUGE-2	ROUGE-L
GPT-3.5	54.33	82.13	65.89	80.79
GPT-4	57.82	85.10	69.55	84.09
ConfigAid	72.51	87.95	80.20	86.68

ConfigAid的配置生成指标评测

大语言模型合成全网配置

Configuration Copilot: From Natural Language to Network-wide Configurations.

研究动机

- 借助大语言模型强大的意图理解能力和文本生成能力，可以极大地减轻网络运维人员的负担。

研究目标

- 解决通用大语言模型生成全网配置文件的严重幻觉问题，如语法错误，语义错误和拓扑错误。

```
cisco Copy code

router bgp 45000
  bgp log-neighbor-changes

  template peer-policy CUSTOMER-A
    route-map SET-COMMUNITY in
    filter-list 20 in

neighbor CUSTOMER-A peer-policy CUSTOMER-A inherit
neighbor CUSTOMER-A peer-policy PRIMARY-IN priority 20
neighbor CUSTOMER-A peer-policy GLOBAL priority 10
```

ChatGPT生成配置

```
router bgp 45000
  template peer-policy CUSTOMER-A
    route-map SET-COMMUNITY in
    filter-list 20 in
    inherit peer-policy PRIMARY-IN 20
    inherit peer-policy GLOBAL 10
```

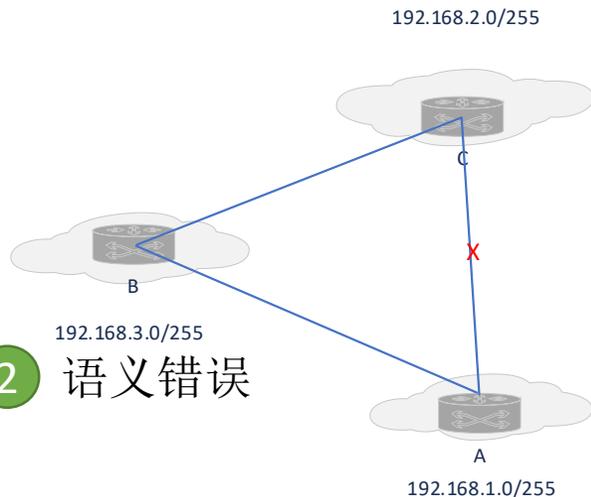
正确配置

1 语法错误

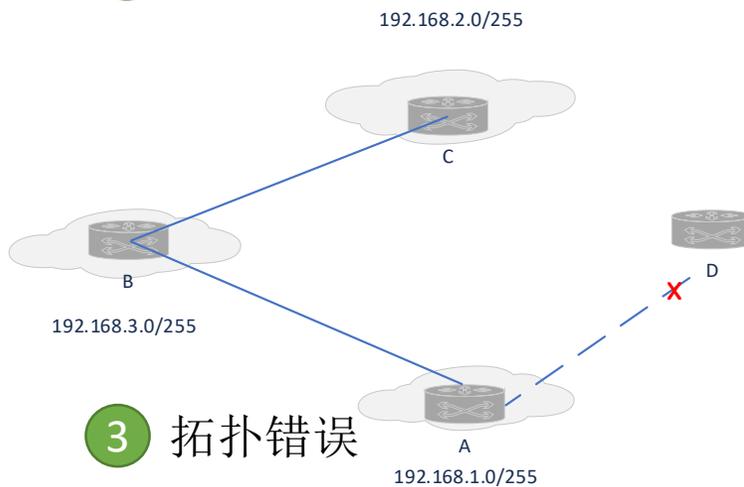
工作流程

- 通过增量预训练，有监督微调，奖励模型建模，强化学习四部分，实现LLM的配置领域适应。

2 语义错误



3 拓扑错误



总 结

1 配置综合概述

2 配置综合实践

3 配 置 翻 译

4 配置大模型





東北大學
Northeastern University

谢谢大家!



中国教育和科研计算机网CERNET · 第二十八/二十九届学术年会