



(12) 发明专利

(10) 授权公告号 CN 108900320 B

(45) 授权公告日 2021.02.09

(21) 申请号 201810565884.3

(22) 申请日 2018.06.04

(65) 同一申请的已公布的文献号  
申请公布号 CN 108900320 A

(43) 申请公布日 2018.11.27

(73) 专利权人 佛山科学技术学院  
地址 528000 广东省佛山市南海区狮山镇  
仙溪水库西路佛山科学技术学院

(72) 发明人 焦波 石建迈 张文生 邢立宁  
戎海武 何敏藩 于辉 王向东

(74) 专利代理机构 广州嘉权专利商标事务所有  
限公司 44205  
代理人 王国标

(51) Int. Cl.  
H04L 12/24 (2006.01)

(56) 对比文件

CN 106462990 A, 2017.02.22

CN 104521160 A, 2015.04.15

US 2015381334 A1, 2015.12.31

US 2014281026 A1, 2014.09.18

jinzhou Zeng, Fangyuan Li. Distributed event-triggered bipartite consensus for multiple agents over signed graph topology. 《2015 34th Chinese Control Conference》. 2015,

审查员 李嵩

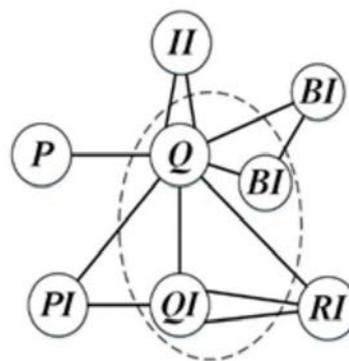
权利要求书14页 说明书28页 附图3页

(54) 发明名称

一种互联网测试床拓扑结构大比例规模缩减方法及装置

(57) 摘要

本发明公开了一种互联网测试床拓扑结构大比例规模缩减方法及装置,通过拓扑二分图结构的分解、采样与合并技术,在保持重要拓扑属性稳定性的条件下,实现自治系统级互联网测试床96%以上的节点删减。本发明可以在保持节点度分布、聚类系数、RichClub连通性和路径长度等重要拓扑属性稳定,即近似保持不变的前提下,大比例地缩减互联网测试床的自治系统级拓扑规模,规模缩减率可以达到96%以上。



1. 一种互联网测试床拓扑结构大比例规模缩减方法,其特征在于,所述方法包括如下步骤:

S1,获取真实世界互联网拓扑结构的探测数据,以简单无向图 $G=(V,E)$ 建模自治系统级互联网拓扑结构,其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域,集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径;

S2,采用互联网拓扑结构模型分解互联网拓扑图 $G=(V,E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $Q_{I_1-RI}$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ ,一个匹配图 $BI-BI$ 和一个内核图 $Core$ ;

S3,调用二分图采样方法1对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $Q_{I_1-RI}$ 进行采样,从而抽取部分节点和边构成采样子图;

其中,所述调用二分图采样方法1包括如下步骤:

S301,输入二分图 $G=(V_1,V_2,E)$ ,其中 $V_1$ 和 $V_2$ 为两个节点集、 $E$ 为边集,参数 $T$ 、 $R$ ,规模缩减率 $R_R$ ;转到步骤S302;

S302,将 $V_2$ 中全部节点按度从大到小顺利排列,并抽取前 $n_h$ 个最大度节点构成节点子集 $V_2^{LH}$ ,其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ;转到步骤S303;

S303,计算 $V_2^{LL} = V_2 / V_2^{LH}$ ,并将二分图 $G$ 分解为 $G^{LH}=(V_1, V_2^{LH}, E^{LH})$ 和 $G^{LL}=(V_1, V_2^{LL}, E^{LL})$ ,其中 $E^{LH} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{LH}\}$ 且 $E^{LL} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ;转到步骤S304;

S304,设 $d_1, d_2, \dots, d_s$ 为图 $G$ 中节点集 $V_1$ 的全部节点度,并设 $d_{max}$ 为节点集 $V_2^{LL}$ 中的最大节点度,计算节点集 $V_2^{LL}$ 的度分布 $\{g(d) | d=1, 2, \dots, d_{max}\}$ ,其中 $g(d)$ 表示 $V_2^{LL}$ 中度为 $d$ 的节点比率,初始化二分图 $G'=(V'_1, V'_2, E') \leftarrow G=(V_1, V_2, E)$ ,初始化二分图 $G^{LH'}=(V'_1, V_2^{LH'}, E^{LH'}) \leftarrow G^{LH}=(V_1, V_2^{LH}, E^{LH})$ ,初始化二分图 $G^{LL'}=(V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL}=(V_1, V_2^{LL}, E^{LL})$ ,初始化二分图 $G^{LL'}$ 中节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) | d=1, 2, \dots, d_{max}\} \leftarrow \{g(d) | d=1, 2, \dots, d_{max}\}$ ,计算需删除边的总数 $E_d = ||E|| \cdot R_R$ ,初始化变量 $t=0$ ;转到步骤S305;

S305,如果 $t < E_d$ ,转到步骤S306;否则二分图 $G'=(V'_1, V'_2, E')$ ,其中 $V'_2 = V_2^{LH'} \cup V_2^{LL'}$ 且 $E' = E^{LH'} \cup E^{LL'}$ ,为最终输出结果,转到步骤S311;

S306,如果 $|S| = \{v | v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} > 0$ ,任意选择集合 $S$ 中一个节点 $v$ ,并计算两个可选边集 $E_o^{LH} = \{(v, v_2) \in E^{LH'} | v_2 \in V_2^{LH'}\}$ 和 $E_o^{LL} = \{(v, v_2) \in E^{LL'} | v_2 \in V_2^{LL'}\}$ ,如果 $|E_o^{LH}| = 0$ ,转到步骤S308;如果 $|E_o^{LL}| = 0$ ,转到步骤S309;否则 $|E_o^{LH}| > 0$ 且 $|E_o^{LL}| > 0$ ,以概率 $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$ 转到步骤S308,并以概率 $1 - P^{LL'}$ 转到步骤S309,其中, $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ , $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ , $m^{LL} = \frac{E_d \cdot ||E^{LL}||}{||E^{LH} \cup E^{LL}||} \cdot (1 + R)$ , $m^{LH} = E_d - m^{LL}$ ;转到步骤S307;

S307,如果 $S = \Phi$ ,以式 $p_k = \sum_{i=1}^k f(d_i)$ 定义的离散概率分布 $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k=1, 2, \dots, s$ )随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ ,其中 $d_1 > d_2 > \dots > d_s > 0$ , $f(d_1), f(d_2), \dots, f(d_s)$ 表示

节点集 $V_1$ 中度为 $d_k$  ( $1 \leq k \leq s$ ) 的节点比率; 计算两个可选边集

$$E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V_1' \wedge v_2 \in V_2^{LH'} \wedge d_v = d\}, \quad E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V_1' \wedge v_2 \in V_2^{LL'} \wedge d_v = d\},$$

如果 $\|E_o^{LH}\| = 0$ , 转到步骤S308; 如果 $\|E_o^{LL}\| = 0$ , 转到步骤S309, 否则, 以概率 $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤S308, 并以概率 $1 - P^{LL'}$  转到步骤S309;

S308, 假设 $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中 $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且 $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ , 计算 $d^h = \arg \max_{d \in D} \{g'(d) - g(d)\}$ , 其中 $D$ 为集合 $V_2^{LL'}$  中节点 $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合, 并计算 $E_o^{LL}$  的子集 $E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} = d^h\}$ , 从 $E_a$  中均等概率分布地采样一条边 $e$ , 并将该边从二分图 $G^{LL'} = (V_1', V_2^{LL'}, E^{LL'})$  和 $G' = (V_1', V_2', E')$  中删除, 即更新 $E^{LL'} \leftarrow E^{LL'} / \{e\}$  且 $E' \leftarrow E' / \{e\}$ , 更新 $G^{LL'}$  中 $V_2^{LL'}$  的节点度分布 $\{g'(d) \mid d = 1, 2, \dots, d_{\max}\}$ , 其中 $g'(d)$  表示 $V_2^{LL'}$  中度为 $d$  的节点比率, 转到步骤S310;

S309, 从 $E_o^{LH}$  中均等概率分布地采样一条边 $e$ , 并将该边从二分图 $G^{LH'} = (V_1', V_2^{LH'}, E^{LH'})$  和 $G' = (V_1', V_2', E')$  中删除, 即更新 $E^{LH'} \leftarrow E^{LH'} / \{e\}$  且 $E' \leftarrow E' / \{e\}$ , 转到步骤S310;

S310, 更新 $t \leftarrow t + 1$ , 转到步骤S305;

S311, 输出: 采样子图 $G' = (V_1', V_2', E')$ ;

S4, 调用匹配图采样方法对匹配图BI-BI进行采样, 抽取部分节点和边构成采样子图;

其中, 所述调用匹配图采样方法包括如下步骤:

S401, 输入二分图 $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$  和匹配图 $G_{BI-BI} = (BI, E_{BI-BI})$ , 采样二分图 $G'_{Q_2-BI} = (BI', Q_2', E'_{Q_2-BI})$ ;

S402, 对每条边 $e = (u, v) \in E_{BI-BI}$ , 其中 $u, v \in BI$ , 生成一个节点度对 $(d_u^{BI}, d_v^{BI})$  (不失一般性, 假设 $d_u^{BI} \leq d_v^{BI}$ ), 并将节点度对 $(d_u^{BI}, d_v^{BI})$  映射为 $x = d_v^{BI} (d_v^{BI} - 1) / 2 + d_u^{BI}$ , 计算集合 $X = \{x = d_v^{BI} (d_v^{BI} - 1) / 2 + d_u^{BI} \mid (u, v) \in E_{BI-BI}\}$ , 并计算节点度对分布函数 $\rho(x) = |\{y \mid y \in X \wedge y = x\}| / |X|$ , 其表示集合 $X$  中出现元素 $x$  的概率值;

S403, 设采样二分图 $G'_{Q_2-BI}$  中 $BI'$  节点的最大度为 $d_m$ , 初始化边集 $E'_{BI-BI} = \Phi$ , 将节点度 $d_1$  以步长1从1增加至 $d_m - 1$ , 并将节点度 $d_2$  以步长1从 $d_1$  增加至 $d_m$ : 对每个节点度对 $(d_1, d_2)$ , 首先计算其期望数量 $m_{d_1, d_2}$  和最大可能数量 $m'_{d_1, d_2}$ , 其中

$$m_{d_1, d_2} = \frac{\|BI'\|}{2} \cdot \rho\left(\frac{d_2(d_2 - 1)}{2} + d_1\right);$$

$$m'_{d_1, d_2} = \begin{cases} \max(\|\{v \in BI' \mid d_v^{BI'} = d_1\}\|, \|\{v \in BI' \mid d_v^{BI'} = d_2\}\|) & \text{如果 } d_1 < d_2 \\ \lfloor \|\{v \in BI' \mid d_v^{BI'} = d_1\}\| / 2 \rfloor & \text{如果 } d_1 = d_2 \end{cases};$$

然后,标准分布概率均等地随机生成  $s = \min(m_{d_1, d_2}, m'_{d_1, d_2})$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, s$ ), 要求对  $\forall 1 \leq i, j \leq s$  满足  $u_i \in BI' \wedge v_i \in BI' \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j \wedge d_{u_i}^{BI'} = d_1 \wedge d_{v_i}^{BI'} = d_2$ ; 最后,更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1,2,\dots,s}$ ;

S404, 对于节点集  $Y = \{v \in BI' \mid \forall (u, w) \in E'_{BI-BI}, u \neq v \wedge w \neq v\}$ , 标准分布地随机生成  $t = \lfloor |BI'| \rfloor / 2 - \lfloor |E'_{BI-BI}| \rfloor$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, t$ ), 要求对  $\forall 1 \leq i, j \leq t$  满足  $u_i \in Y \wedge v_i \in Y \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j$ , 并更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1,2,\dots,t}$ , 此时,生成的采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$  为本节的输出结果;

S405, 采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$ ;

S5, 调用二分图采样方法2对二分图  $Q_4$ - $RI_1$  进行采样, 抽取部分节点和边构成采样子图; 其中, 所述调用二分图采样方法2包括如下步骤:

S501, 输入二分图  $G_{Q_4-RI_1} = (RI_1, Q_4, E_{Q_4-RI_1})$ 、采样二分图  $G'_{Q_4-RI_1} = (RI'_1, Q'_4, E'_{Q_4-RI_1})$ ;

S502, 计算节点集  $RI'_1 = \{w \in RI'_1 \mid \exists v \in Q_4, (w, v) \in E_{Q_4-RI_1}\}$ 、节点集  $Q'_4 = \{v \in Q_4 \mid \exists w \in RI'_1, (w, v) \in E_{Q_4-RI_1}\}$  和边集  $E'_{Q_4-RI_1} = \{(w, v) \in E_{Q_4-RI_1} \mid w \in RI'_1 \wedge v \in Q'_4\}$ ;

S503, 由节点集  $RI'_1$ 、节点集  $Q'_4$  和边集  $E'_{Q_4-RI_1}$  生成的二分图  $G'_{Q_4-RI_1} = (RI'_1, Q'_4, E'_{Q_4-RI_1})$ , 即为最终的输出结果;

S504, 输出采样二分图  $G'_{Q_4-RI_1} = (RI'_1, Q'_4, E'_{Q_4-RI_1})$ ;

S6, 调用二分图采样方法3对二分图  $QI$ - $PI$  和二分图  $Q$ - $P$  进行采样, 抽取部分节点和边构成采样子图;

其中, 所述调用二分图采样方法3包括如下步骤:

S601, 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1$  和  $V_2$  为两个节点集、 $E$  为边集、节点集  $V_2$  包含节点的度均为1, 节点集  $\bar{V}_1 \subseteq V_1$ , 参数  $T$ , 规模缩减率  $R_R$ ; 转到步骤S602;

S602, 将  $V_1$  中全部节点按度从大到小顺利排列, 并抽取前  $n_h$  个最大度节点构成节点子集  $V_1^H$ , 其中这  $n_h$  个节点中相同度重复数的最大值为  $T$ ; 转到步骤S603;

S603, 计算  $V_1^L = V_1 / V_1^H$ , 设  $d_1, d_2, \dots, d_s$  为二分图  $G$  中节点集  $V_1$  的全部节点度, 初始化二分图  $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ , 初始化节点集  $V_1^{L'} \leftarrow V_1^L$ , 初始化节点集  $V_1^{H'} \leftarrow V_1^H$ , 计算最终需要删除边的总数  $E_d = \lfloor |E| \rfloor \cdot R_R$ , 初始化变量  $t = 0$ ; 转到步骤S604;

S604, 如果  $t < E_d$ , 则以概率  $P = \gamma^{L'} / (\gamma^{H'} + \gamma^{L'})$  转到步骤S605, 并以概率  $1 - P$  转到步骤S608; 否则,  $G' = (V'_1, V'_2, E')$ , 其中  $V'_1 = V_1^{H'} \cup V_1^{L'}$ , 为最终输出结果, 转到步骤S610; 其中,

$$\gamma^{L'} = m^L - \left( \sum_{v \in V_1^L} d_v - \sum_{v \in V_1^{L'}} d_v \right), \quad \gamma^{H'} = m^H - \left( \sum_{v \in V_1^H} d_v - \sum_{v \in V_1^{H'}} d_v \right),$$

$$m^L = \frac{E_d \cdot \sum_{v \in V_1^L} d_v}{\sum_{v \in V_1^{H'} \cup V_1^{L'}} d_v} \quad \text{且 } m^H = E_d - m^L;$$

S605, 如果  $\left\| S = \left\{ v \mid v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\} \right\} / \bar{V}_1 \right\| > 0$ , 其中  $\bar{V}_1 = \{v \mid d_v = 1 \wedge v \in \bar{V}_1\}$ , 则随机选择集合  $S$  中一个节点  $v$ , 转到步骤S607, 否则, 转到步骤S606;

S606, 以式  $p_k = \sum_{i=1}^k f(d_i)$  定义的离散概率分布  $p_k / \sum_{i=1,2,\dots,s} p_i$  ( $k=1, 2, \dots, s$ ), 随机采样一个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中,  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  在二分图  $G = (V_1, V_2, E)$  中度为  $d_k$  ( $1 \leq k \leq s$ ) 的节点比率, 计算两个节点集  $V^1 = \{v | v \in V_1^{L'} \cap \bar{V}_1 \wedge d_v = d\}$  和  $V^2 = \{v | v \in V_1^{L'} / \bar{V}_1 \wedge d_v = d\}$ , 如果  $||V^2|| > 0$ , 则随机选择集合  $V^2$  中一个节点  $v$ , 否则随机选择集合  $V^1$  中一个节点  $v$ , 转到步骤S607;

S607, 如果  $d_v \geq 2 \vee v \notin \bar{V}_1$ , 则随机选择一条与节点  $v$  相邻的边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E')$ ,  $E' \leftarrow E' / \{e\}$ , 如果  $d_v = 0$ , 则更新  $V_1^{L'} \leftarrow V_1^{L'} / \{v\}$ , 如果  $d_v < 2 \wedge v \in \bar{V}_1$ , 则此步不作任何操作; 转到步骤S609;

S608, 从集合  $\{(w, v) \in E' | w \in V_2' \wedge v \in V_1^{H'}\} / \bar{V}_1$  中随机选择一条边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E')$ :  $E' \leftarrow E' / \{e\}$ ; 如果  $d_v = 0$ , 则更新  $V_1^{H'} \leftarrow V_1^{H'} / \{v\}$ ; 转到步骤S609;

S609, 更新  $t \leftarrow t+1$ , 转到步骤S604;

S610, 输出: 采样子图  $G' = (V'_1, V'_2, E')$ ; 其中, 输入的节点集  $\bar{V}_1$  被包含于采样子图  $G' = (V'_1, V'_2, E')$ , 即  $\bar{V}_1 \subseteq V_1'$ ; 图  $G$  中节点集  $V_2$  包含节点的度均为1, 图  $G'$  中节点集  $V_2'$  包含节点的度均为1;

S7, 调用二分图采样方法4对二分图  $Q_3$ -PI 进行采样, 抽取部分节点和边构成采样子图;

其中, 所述调用二分图采样方法4包括如下步骤:

S701, 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1$  和  $V_2$  为两个节点集、 $E$  为边集, 多对一映射函数  $\psi: V_1 \rightarrow W$ , 设  $\psi(v_1)$  为节点集  $W$  中与节点  $v_1 \in V_1$  对应的唯一节点, 对任意节点  $w \in W$  预选设定参数  $L(w)$ , 预选给定节点度分布  $\{g(d) | d=1, 2, \dots\}$ , 预先设定参数  $T, R$ , 规模缩减率  $R_R$ ; 转到步骤S702;

S702, 将  $V_2$  中全部节点按度从大到小顺利排列, 并抽取前  $n_h$  个最大度节点构成节点子集  $V_2^{LH}$ , 其中这  $n_h$  个节点中相同度重复数的最大值为  $T$ ; 转到步骤S703;

S703, 计算  $V_2^{LL} = V_2 / V_2^{LH}$ , 并将二分图  $G$  分解为  $G^{LH} = (V_1, V_2^{LH}, E^{LH})$  和  $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 其中  $E^{LH} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{LH}\}$  且  $E^{LL} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ; 转到步骤S704;

S704, 设  $d_1, d_2, \dots, d_s$  为二分图  $G$  中节点集  $V_1$  的全部节点度, 并设  $d_{\max}$  为节点集  $V_2^{LL}$  中的最大节点度, 初始化二分图  $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ , 初始化二分图  $G^{LH'} = (V_1', V_2^{LH'}, E^{LH'}) \leftarrow G^{LH} = (V_1, V_2^{LH}, E^{LH})$ , 初始化二分图  $G^{LL'} = (V_1', V_2^{LL'}, E^{LL'}) \leftarrow G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 计算二分图  $G^{LL'}$  中节点集  $V_2^{LL'}$  的度分布  $\{g'(d) | d=1, 2, \dots, d_{\max}\}$ , 其中,  $g'(d)$  表示  $V_2^{LL'}$  中度为  $d$  的节点比率, 计算最终需要删除边的总数  $E_d = ||E|| \cdot R_R$ , 其中, 初始化变量  $t=0, \eta=0$ ; 转到步骤S705;

S705, 如果  $t < E_d$ , 转到步骤S706; 否则二分图  $G' = (V'_1, V'_2, E')$ , 其中  $V_2' = V_2^{LH'} \cup V_2^{LL'}$  且  $E' = E^{LH'} \cup E^{LL'}$ , 为最终输出结果, 转到步骤S713;

S706, 如果  $\eta > 0 \wedge \|\tilde{V}_1\| > 0$ , 则更新  $\eta \leftarrow \eta - 1$ , 计算两个可选边集

$$E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\} \text{ 和 } E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\},$$

$$\text{其中 } \tilde{V}_1 = \left\{ v_1 \in V_1' \mid \frac{\|N'(\psi(v_1))\|}{L(\psi(v_1))} > 1 \wedge d_v = 1 \right\}, \text{ 且集合 } N'(\psi(v_1)) = \{v \mid v \in V_1' \wedge \psi(v)$$

$= \psi(v_1)\}$ ,

并转到步骤S709, 否则, 转到步骤S707;

S707, 如果  $\|S = \{v \mid v \in V_1' \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}\| > 0$ , 选择集合S中一个节点v, 设该节点的度为  $d = d_v$ ; 否则以式  $p_k = \sum_{i=1}^k f(d_i)$  定义的离散概率分布  $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k = 1, 2, \dots, s$ ) 随机采样一个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  中度为  $d_k$  ( $1 \leq k \leq s$ ) 的节点比率, 转到步骤S708;

S708, 如果  $d = 1 \wedge \|\tilde{V}_1\| = 0$ , 则更新  $\eta \leftarrow \eta + 1$ , 转到步骤S712; 如果  $d = 1 \wedge \|\tilde{V}_1\| > 0$ , 则计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ , 转到步骤S709; 如果  $d > 1$ , 则计算两个可选边集

$$E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V_1' \wedge v_2 \in V_2^{LH'} \wedge d_v = d\} \text{ 和 } E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V_1' \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}, \text{ 转到步骤S709;}$$

S709, 如果  $\|E_o^{LH}\| = 0$ , 转到步骤S710; 如果  $\|E_o^{LL}\| = 0$ , 转到步骤S711; 如果  $\|E_o^{LH}\| > 0 \wedge \|E_o^{LL}\| > 0$ , 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤S710, 以概率  $1 - P^{LL'}$  转到步骤S711,

$$\text{其中, } \gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||), \gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||),$$

$$m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R), m^{LH} = E_d \cdot m^{LL};$$

S710, 假设  $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ , 计算  $D^h = \{d \in D \mid g'(d) - g(d) > 1 / \|V_2^{LL'}\|\}$ , 其中D为集合  $V_2^{LL'}$  中节点  $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合; 如果  $||D^h|| = 0$ , 则更新  $D^h \leftarrow \{\arg \max_{d \in D} \{g'(d) - g(d)\}\}$ ;

$$\text{首先, 计算集合 } E_o^{LL} \text{ 的一个子集: } E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} \in D^h\},$$

假设  $E_a = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_q^l, v_q^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_q^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_q^h \in V_2^{LL'}$ ; 计算  $\psi' = \arg \max_{v \in V} \{||N'(\psi(v))|| / L(\psi(v))\}$ , 其中  $V = \{v_1^l, v_2^l, \dots, v_q^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$ ,

然后,计算集合 $E_a$ 的一个子集, $E_b = \{(v_i^l, v_i^r) \in E_a \mid \psi(v_i^l) = \psi^l\}$ ,

最后,从集合 $E_b$ 中随机选择一条边 $e$ ,并将该边从二分图 $G^{LL'} = (V_1', V_2^{LL'}, E^{LL'})$ 和 $G' = (V_1', V_2', E')$ 中删除,即更新 $E^{LL'} \leftarrow E^{LL'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,更新二分 $G^{LL'}$ 中 $V_2^{LL'}$ 的节点度分布 $\{g'(d) \mid d=1, 2, \dots, d_{\max}\}$ ,其中 $g'(d)$ 表示 $V_2^{LL'}$ 中度为 $d$ 的节点比率,转到步骤S712;

S711,假设 $E_o^{LH} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ ,其中 $v_1^l, v_2^l, \dots, v_r^l \in V_1'$ 且 $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LH'}$ ;计算 $\psi^l = \arg \max_{\psi(v)} \{ ||N'(\psi(v))|| / L(\psi(v)) \}$ ,其中 $V = \{v_1^l, v_2^l, \dots, v_r^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$ ,

计算集合 $E_o^{LH}$ 的一个子集: $E_b = \{(v_i^l, v_i^h) \in E_o^{LH} \mid \psi(v_i^l) = \psi^l\}$ ,

并从集合 $E_b$ 中随机选择一条边 $e$ ,并将该边从二分图 $G^{LH'} = (V_1', V_2^{LH'}, E^{LH'})$ 和 $G' = (V_1', V_2', E')$ 中删除,即更新 $E^{LH'} \leftarrow E^{LH'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,转到步骤S712;

S712,更新 $t \leftarrow t+1$ ,转到步骤S705;

S713,输出:采样子图 $G' = (V_1', V_2', E')$ ;其中,对 $\forall w \in W$ ,集合 $N'(w) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = w\}$ ,其中 $W$ 为步骤S701的输入, $V_1' \subseteq V_1$ ,S701至S713期望对 $\forall w \in W$ 满足 $||N'(w)||$ 接近于 $L(w)$ ;

S8,对S3至S7得到的所有采样子图进行合并,得到这些采样子图的合并图;

S9,调用内核图采样方法对内核图Core进行采样,从而抽取部分节点和边构成采样子图;

其中,所述调用内核图采样方法包括如下步骤:

S901,输入原始互联网拓扑 $G = (V, E)$ 、合并子图 $G'_m$ 和内核子图 $G'_{Core} = (V'_{Core}, E'_{Core})$ ;

S902,计算原始拓扑 $G$ 的平均节点度 $d_{avg}$ ,计算合并子图 $G'_m$ 的边数 $m(G'_m)$ 和节点数 $n(G'_m)$ ,将所有边 $(w, v) \in E'_{Core}$ 按照 $\max(d_w(G'_{Core}), d_v(G'_{Core}))$ 不降顺序排列;

S903,从边集 $E'_{Core}$ 中删除排在前面的 $||E'_{Core}|| - d_{avg} \cdot n(G'_m) / 2 + m(G'_m)$ 条边,并将剩余边组成的集合定义为 $E''_{Core}$ ;

S904, $E''_{Core}$ 为最终输出的内核边集;

S905,输出内核边集 $E''_{Core}$ ;

S10,对S8得到的合并图和S9得到的采样子图进行合并,从而得到真实世界互联网拓扑图 $G = (V, E)$ 的采样子图;

S11,依据S10得到的真实世界互联网拓扑图的规模缩减采样子图,配置互联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境;

其中,所述互联网拓扑结构模型分解为八类节点集,其中节点集OI表示系统噪声;

$$\left\{ \begin{array}{l} P = \{v \in V \mid d_v = 1\} \\ Q = \{v \in V \mid \exists w, (v, w) \in E, w \in P\} \\ PI = \{v \in V_I \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) > 1\} \\ QI = \{v \in V_I \mid \exists w, (v, w) \in E_I, w \in PI\} \\ RI = \{v \in V_I \mid d_I(v) \geq 2 \wedge \forall (v, w) \in E_I, w \in QI\} \\ BI = \{v \in V_I \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) = 1\} \\ II = \{v \in V_I \mid d_I(v) = 0\} \\ OI = V_I / (PI \cup QI \cup RI \cup BI \cup II) \end{array} \right.$$

其中,  $V$ 和 $E$ 分别为互联网拓扑图 $G = (V, E)$ 的节点集和边集,  $v$ 为拓扑图 $G = (V, E)$ 中的节点,  $d_v$ 为拓扑图 $G = (V, E)$ 中节点 $v$ 的度;  $G_I = (V_I, E_I)$ 为拓扑图 $G = (V, E)$ 中由节点集 $R(G) = V / (P \cup Q)$ 和两个端点均属于 $R(G)$ 的所有边构成边集生成的子图;  $V_I, E_I$ :分别为子图 $G_I = (V_I, E_I)$ 的节点集和边集;  $d_I(v)$ 为子图 $G_I = (V_I, E_I)$ 中节点 $v$ 的度。

2. 根据权利要求1所述的一种互联网测试床拓扑结构大比例规模缩减方法, 其特征在于, 互联网拓扑结构模型的八类节点集节点之间的连接关系为: 每个 $P$ 节点仅且必被连接至唯一的 $Q$ 节点, 每个 $II$ 节点仅且必被连接至至少两个 $Q$ 节点, 每个 $PI$ 节点仅且必被连接至唯一的 $QI$ 节点和至少一个 $Q$ 节点, 每个 $BI$ 节点仅且必被连接至另外一个 $BI$ 节点和至少一个 $Q$ 节点, 每个 $RI$ 节点仅且必被连接至至少两个 $QI$ 节点, 此外 $QI$ 和 $RI$ 节点可能被连接至 $Q$ 节点, 由连接关系将自治系统级互联网拓扑分解为七个二分图 $Q_1-II, Q-P, Q_2-BI, Q_3-PI, QI-PI, QI_1-RI$ 和 $Q_4-RI_1$ , 其中节点集

$$\begin{array}{ll} Q_1 = \{v \in Q \mid \exists w \in II, (v, w) \in E\} & , \quad Q_2 = \{v \in Q \mid \exists w \in BI, (v, w) \in E\} \\ Q_3 = \{v \in Q \mid \exists w \in PI, (v, w) \in E\} & , \quad Q_4 = \{v \in Q \mid \exists w \in RI, (v, w) \in E\} \end{array}$$

$QI_1 = \{v \in QI \mid \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI \mid \exists v \in Q, (v, w) \in E\}$ , 一个匹配图 $BI-BI$ 和一个内核图 $Core$ ,

其中, 二分图是指图中全部节点可以被划分至两个不同的节点集, 且图中每条边的两个端点必定分别属于这两个不同的节点集, 二分图 $Q_1-II$ 是指由节点集 $Q_1$ 、节点集 $II$ 和两个端点分别属于 $Q_1$ 和 $II$ 的所有边构成边集生成的二分图结构, 二分图 $Q-P, Q_2-BI, Q_3-PI, QI-PI, QI_1-RI$ 和 $Q_4-RI_1$ 的物理意义解释方法与二分图 $Q_1-II$ 相同; 匹配图 $BI-BI$ 是由节点集 $BI$ 和两个端点都属于 $BI$ 的所有边构成边集生成的图结构; 内核图 $Core$ 是由节点集 $Q \cup QI$ 和连接该节点集内部节点的所有边构成边集生成的图结构。

3. 根据权利要求1所述的一种互联网测试床拓扑结构大比例规模缩减方法, 其特征在于, 在 $S_2$ 中, 互联网拓扑图分解方法包括如下步骤:

S201, 读取互联网拓扑图 $G = (V, E)$ 的七个节点集 $P, Q, II, BI, QI, PI$ 和 $RI$ ;

S202, 计算二分图 $Q_1-II$ :  $G_{Q_1-II} = (II, Q_1, E_{Q_1-II})$ , 其中  $Q_1 = \{v \in Q \mid \exists w \in II, (v, w) \in E\}$  且  $E_{Q_1-II} = \{e = (v, w) \mid e \in E \wedge v \in II \wedge w \in Q_1\}$ ;

S203, 计算二分图 $Q_2-BI$ :  $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$ , 其中  $Q_2 = \{v \in Q \mid \exists w \in BI, (v, w) \in E\}$  且  $E_{Q_2-BI} = \{e = (v, w) \mid e \in E \wedge v \in BI \wedge w \in Q_2\}$ ;

S204, 计算二分图 $QI_1-RI: G_{QI_1-RI} = (RI, QI_1, E_{QI_1-RI})$ , 其中 $QI_1 = \{v \in QI \mid \exists w \in RI, (v, w) \in E\}$ 且 $E_{QI_1-RI} = \{e = (v, w) \mid e \in E \wedge v \in RI \wedge w \in QI_1\}$ ;

S205, 计算二分图 $Q_4-RI_1: G_{Q_4-RI_1} = (RI_1, Q_4, E_{Q_4-RI_1})$ , 其中 $Q_4 = \{v \in Q \mid \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI \mid \exists v \in Q, (v, w) \in E\}$ 且 $E_{Q_4-RI_1} = \{e = (v, w) \mid e \in E \wedge v \in RI_1 \wedge w \in Q_4\}$ ;

S206, 计算二分图 $QI-PI: G_{QI-PI} = (QI, PI, E_{QI-PI})$ , 其中 $E_{QI-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in QI\}$ ;

S207, 计算二分图 $Q_3-PI: G_{Q_3-PI} = (PI, Q_3, E_{Q_3-PI})$ , 其中 $Q_3 = \{v \in Q \mid \exists w \in PI, (v, w) \in E\}$ 且 $E_{Q_3-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in Q_3\}$ ;

S208, 计算二分图 $Q-P: G_{Q-P} = (Q, P, E_{Q-P})$ , 其中 $E_{Q-P} = \{e = (v, w) \mid e \in E \wedge v \in P \wedge w \in Q\}$ ;

S209, 计算匹配图 $BI-BI: G_{BI-BI} = (BI, E_{BI-BI})$ , 其中 $E_{BI-BI} = \{e = (v, w) \mid e \in E \wedge v, w \in BI\}$ ;

S210, 计算内核图 $Core: G_{Core} = (V_{Core}, E_{Core})$ , 其中 $V_{Core} = Q \cup QI$ 且 $E_{Core} = \{e = (v, w) \mid e \in E \wedge v, w \in V_{Core}\}$ ;

S211, 输出七个二分图 $G_{Q_1-II}$ 、 $G_{Q_2-BI}$ 、 $G_{QI_1-RI}$ 、 $G_{Q_4-RI_1}$ 、 $G_{QI-PI}$ 、 $G_{Q_3-PI}$ 、 $G_{Q-P}$ , 一个匹配图 $G_{BI-BI}$ 和一个内核图 $G_{Core}$ ,

其中, $\in$ 表示“属于”; $\exists$ 表示“存在”; $\wedge$ 表示“逻辑与”; $\cup$ 表示集合的并集; $V$ 和 $E$ 分别表示互联网拓扑图 $G = (V, E)$ 的节点集和边集; $P$ 、 $Q$ 、 $II$ 、 $BI$ 、 $QI$ 、 $PI$ 和 $RI$ 分别为互联网拓扑结构模型分解得到的七类节点集。

4. 一种互联网测试床拓扑结构大比例规模缩减装置, 其特征在于, 所述装置包括:

建模单元, 用于获取真实世界互联网拓扑结构的探测数据, 以简单无向图 $G = (V, E)$ 建模自治系统级互联网拓扑结构, 其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域, 集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径;

分解拓扑单元, 用于采用互联网拓扑结构模型分解互联网拓扑图 $G = (V, E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $QI_1-RI$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ , 一个匹配图 $BI-BI$ 和一个内核图 $Core$ ;

二分图采样1单元, 用于调用二分图采样方法1对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $QI_1-RI$ 进行采样, 从而抽取部分节点和边构成采样子图;

其中, 所述调用二分图采样方法1包括如下步骤:

S301, 输入二分图 $G = (V_1, V_2, E)$ , 其中 $V_1$ 和 $V_2$ 为两个节点集、 $E$ 为边集, 参数 $T$ 、 $R$ , 规模缩减率 $R_R$ ; 转到步骤S302;

S302, 将 $V_2$ 中全部节点按度从大到小顺利排列, 并抽取前 $n_h$ 个最大度节点构成节点子集 $V_2^{LH}$ , 其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ; 转到步骤S303;

S303, 计算 $V_2^{LL} = V_2 / V_2^{LH}$ , 并将二分图 $G$ 分解为 $G^{LH} = (V_1, V_2^{LH}, E^{LH})$ 和 $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 其中 $E^{LH} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LH}\}$ 且 $E^{LL} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ; 转到步骤S304;

S304, 设 $d_1, d_2, \dots, d_s$ 为图 $G$ 中节点集 $V_1$ 的全部节点度, 并设 $d_{max}$ 为节点集 $V_2^{LL}$ 中的最大节点度, 计算节点集 $V_2^{LL}$ 的度分布 $\{g(d) \mid d = 1, 2, \dots, d_{max}\}$ , 其中 $g(d)$ 表示 $V_2^{LL}$ 中节点度为 $d$ 的节点比

率, 初始化二分图  $G' = (V'_1, V'_2, E')$   $\leftarrow G = (V_1, V_2, E)$ , 初始化二分图  $G^{LH'} = (V'_1, V_2^{LH'}, E^{LH'}) \leftarrow G^{LH} = (V_1, V_2^{LH}, E^{LH})$ , 初始化二分图

$G^{LL'} = (V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 初始化二分图  $G^{LL'}$  中节点集  $V_2^{LL'}$  的度分布  $\{g'(d) \mid d = 1, 2, \dots, d_{\max}\} \leftarrow \{g(d) \mid d = 1, 2, \dots, d_{\max}\}$ , 计算需删除边的总数  $E_d = ||E|| \cdot R_R$ , 初始化变量  $t = 0$ ; 转到步骤 S305;

S305, 如果  $t < E_d$ , 转到步骤 S306; 否则二分图  $G' = (V'_1, V'_2, E')$ , 其中  $V'_2 = V_2^{LH'} \cup V_2^{LL'}$  且  $E' = E^{LH'} \cup E^{LL'}$ , 为最终输出结果, 转到步骤 S311;

S306, 如果  $\|S = \{v \mid v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}\| > 0$ , 任意选择集合  $S$  中一个节点  $v$ , 并计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v_2 \in V_2^{LH'}\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v_2 \in V_2^{LL'}\}$ , 如果  $\|E_o^{LH}\| = 0$ , 转到步骤 S308; 如果  $\|E_o^{LL}\| = 0$ , 转到步骤 S309; 否则  $\|E_o^{LH}\| > 0$  且  $\|E_o^{LL}\| > 0$ , 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤 S308, 并以概率  $1 - P^{LL'}$  转到步骤 S309, 其中,  $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ ,  $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ ,  $m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R)$ ,  $m^{LH} = E_d - m^{LL}$ ; 转到步骤 S307;

S307, 如果  $S = \Phi$ , 以式  $p_k = \sum_{i=1}^k f(d_i)$  定义的离散概率分布  $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k = 1, 2, \dots, s$ ) 随机采样一个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  中度为  $d_k$  ( $1 \leq k \leq s$ ) 的节点比率; 计算两个可选边集

$E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V'_1 \wedge v_2 \in V_2^{LH'} \wedge d_v = d\}$ ,  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V'_1 \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}$ , 如果  $\|E_o^{LH}\| = 0$ , 转到步骤 S308; 如果  $\|E_o^{LL}\| = 0$ , 转到步骤 S309,

否则, 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤 S308, 并以概率  $1 - P^{LL'}$  转到步骤 S309;

S308, 假设  $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V'_1$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ , 计算  $d^h = \arg \max_{d \in D} \{g'(d) - g(d)\}$ , 其中  $D$  为集合  $V_2^{LL'}$  中节点  $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合, 并计算  $E_o^{LL}$  的子集  $E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} = d^h\}$ , 从  $E_a$  中均等概率分布地采样一条边  $e$ , 并将该边从二分图  $G^{LL'} = (V'_1, V_2^{LL'}, E^{LL'})$  和  $G' = (V'_1, V'_2, E')$  中删除, 即更新  $E^{LL'} \leftarrow E^{LL'} / \{e\}$  且  $E' \leftarrow E' / \{e\}$ , 更新  $G^{LL'}$  中  $V_2^{LL'}$  的节点度分布  $\{g'(d) \mid d = 1, 2, \dots, d_{\max}\}$ , 其中  $g'(d)$  表示  $V_2^{LL'}$  中度为  $d$  的节点比率, 转到步骤 S310;

S309, 从  $E_o^{LH}$  中均等概率分布地采样一条边  $e$ , 并将该边从二分图  $G^{LH'} = (V'_1, V_2^{LH'}, E^{LH'})$  和  $G' = (V'_1, V'_2, E')$  中删除, 即更新  $E^{LH'} \leftarrow E^{LH'} / \{e\}$  且  $E' \leftarrow E' / \{e\}$ , 转到步骤 S310;

S310, 更新  $t \leftarrow t + 1$ , 转到步骤 S305;

S311, 输出: 采样子图  $G' = (V'_1, V'_2, E')$ ;

匹配图采样单元, 用于调用匹配图采样方法对匹配图 BI-BI 进行采样, 抽取部分节点和

边构成采样子图；

其中，所述调用匹配图采样方法包括如下步骤：

S401，输入二分图  $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$  和匹配图  $G_{BI-BI} = (BI, E_{BI-BI})$ ，采样二分图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$ ；

S402，对每条边  $e = (u, v) \in E_{BI-BI}$ ，其中  $u, v \in BI$ ，生成一个节点度对  $(d_u^{BI}, d_v^{BI})$ （不失一般性，假设  $d_u^{BI} \leq d_v^{BI}$ ），并将节点度对  $(d_u^{BI}, d_v^{BI})$  映射为  $x = d_v^{BI}(d_v^{BI} - 1)/2 + d_u^{BI}$ ，计算集合  $X = \{x = d_v^{BI}(d_v^{BI} - 1)/2 + d_u^{BI} \mid (u, v) \in E_{BI-BI}\}$ ，并计算节点度对分布函数  $\rho(x) = |\{y \mid y \in X \wedge y = x\}| / |X|$ ，其表示集合  $X$  中出现元素  $x$  的概率值；

S403，设采样二分图  $G'_{Q_2-BI}$  中  $BI'$  节点的最大度为  $d_m$ ，初始化边集  $E'_{BI-BI} = \Phi$ ，将节点度  $d_1$  以步长1从1增加至  $d_m - 1$ ，并将节点度  $d_2$  以步长1从  $d_1$  增加至  $d_m$ ：对每个节点度对  $(d_1, d_2)$ ，首先计算其期望数量  $m_{d_1, d_2}$  和最大可能数量  $m'_{d_1, d_2}$ ，其中

$$m_{d_1, d_2} = \frac{\|BI'\|}{2} \cdot \rho\left(\frac{d_2(d_2 - 1)}{2} + d_1\right);$$

$$m'_{d_1, d_2} = \begin{cases} \max(\|\{v \in BI' \mid d_v^{BI'} = d_1\}\|, \|\{v \in BI' \mid d_v^{BI'} = d_2\}\|) & \text{如果 } d_1 < d_2 \\ \lfloor \|\{v \in BI' \mid d_v^{BI'} = d_1\}\| / 2 \rfloor & \text{如果 } d_1 = d_2 \end{cases};$$

然后，标准分布概率均等地随机生成  $s = \min(m_{d_1, d_2}, m'_{d_1, d_2})$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, s$ )，要求对  $\forall 1 \leq i, j \leq s$  满足  $u_i \in BI' \wedge v_i \in BI' \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j \wedge d_{u_i}^{BI'} = d_1 \wedge d_{v_i}^{BI'} = d_2$ ；最后，更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1, 2, \dots, s}$ ；

S404，对于节点集  $Y = \{v \in BI' \mid \forall (u, w) \in E'_{BI-BI}, u \neq v \wedge w \neq v\}$ ，标准分布地随机生成  $t = \lfloor \|BI'\| / 2 - \|E'_{BI-BI}\| \rfloor$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, t$ )，要求对  $\forall 1 \leq i, j \leq t$  满足  $u_i \in Y \wedge v_i \in Y \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j$ ，并更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1, 2, \dots, t}$ ，此时，生成的采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$  为本节的输出结果；

S405，采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$ ；

二分图采样2单元，用于调用二分图采样方法2对二分图  $Q_4-RI_1$  进行采样，抽取部分节点和边构成采样子图；

其中，所述调用二分图采样方法2包括如下步骤：

S501，输入二分图  $G_{Q_4-RI_1} = (RI_1, Q_4, E_{Q_4-RI_1})$ ，采样二分图  $G'_{Q_4-RI_1} = (RI'_1, Q'_4, E'_{Q_4-RI_1})$ ；

S502，计算节点集  $RI'_1 = \{w \in RI'_1 \mid \exists v \in Q_4, (w, v) \in E_{Q_4-RI_1}\}$ 、节点集  $Q'_4 = \{v \in Q_4 \mid \exists w \in RI'_1, (w, v) \in E_{Q_4-RI_1}\}$  和边集  $E'_{Q_4-RI_1} = \{(w, v) \in E_{Q_4-RI_1} \mid w \in RI'_1 \wedge v \in Q'_4\}$ ；

S503，由节点集  $RI'_1$ 、节点集  $Q'_4$  和边集  $E'_{Q_4-RI_1}$  生成的二分图  $G'_{Q_4-RI_1} = (RI'_1, Q'_4, E'_{Q_4-RI_1})$ ，即为最终的输出结果；

S504, 输出采样二分图  $G'_{Q_4-R_1} = (R'_1, Q'_4, E'_{Q_4-R_1})$ ;

二分图采样3单元, 用于调用二分图采样方法3对二分图QI-PI和二分图Q-P进行采样, 抽取部分节点和边构成采样子图;

其中, 所述调用二分图采样方法3包括如下步骤:

S601, 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1$  和  $V_2$  为两个节点集、 $E$  为边集、节点集  $V_2$  包含节点的度均为1, 节点集  $\bar{V}_1 \subseteq V_1$ , 参数  $T$ , 规模缩减率  $R_R$ ; 转到步骤S602;

S602, 将  $V_1$  中全部节点按度从大到小顺利排列, 并抽取前  $n_h$  个最大度节点构成节点子集  $V_1^H$ , 其中这  $n_h$  个节点中相同度重复数的最大值为  $T$ ; 转到步骤S603;

S603, 计算  $V_1^L = V_1 / V_1^H$ , 设  $d_1, d_2, \dots, d_s$  为二分图  $G$  中节点集  $V_1$  的全部节点度, 初始化二分图  $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ , 初始化节点集  $V_1^{L'} \leftarrow V_1^L$ , 初始化节点集  $V_1^{H'} \leftarrow V_1^H$ , 计算最终需要删除边的总数  $E_d = ||E|| \cdot R_R$ , 初始化变量  $t = 0$ ; 转到步骤S604;

S604, 如果  $t < E_d$ , 则以概率  $P = \gamma^{L'} / (\gamma^{H'} + \gamma^{L'})$  转到步骤S605, 并以概率  $1 - P$  转到步骤S608; 否则,  $G' = (V'_1, V'_2, E')$ , 其中  $V'_1 = V_1^{H'} \cup V_1^{L'}$ , 为最终输出结果, 转到步骤S610; 其中,

$$\gamma^{L'} = m^L - \left( \sum_{v \in V_1^{L'}} d_v - \sum_{v \in V_1^{L'}} d_v \right), \quad \gamma^{H'} = m^H - \left( \sum_{v \in V_1^{H'}} d_v - \sum_{v \in V_1^{H'}} d_v \right),$$

$$m^L = \frac{E_d \cdot \sum_{v \in V_1^{L'}} d_v}{\sum_{v \in V_1^{H'} \cup V_1^{L'}} d_v} \text{ 且 } m^H = E_d - m^L;$$

S605, 如果  $\left| S = \left\{ v \mid v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\} \right\} / \bar{V}_1 \right| > 0$ , 其中  $\bar{V}_1 = \{v \mid d_v = 1 \wedge v \in \bar{V}_1\}$ , 则随机选择集合  $S$  中一个节点  $v$ , 转到步骤S607, 否则, 转到步骤S606;

S606, 以式  $p_k = \sum_{i=1}^k f(d_i)$  定义的离散概率分布  $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k = 1, 2, \dots, s$ ), 随机采样一个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中,  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  在二分图  $G = (V_1, V_2, E)$  中度为  $d_k$  ( $1 \leq k \leq s$ ) 的节点比率, 计算两个节点集  $V^1 = \{v \mid v \in V_1^{L'} \cap \bar{V}_1 \wedge d_v = d\}$  和  $V^2 = \{v \mid v \in V_1^{L'} / \bar{V}_1 \wedge d_v = d\}$ , 如果  $||V^2|| > 0$ , 则随机选择集合  $V^2$  中一个节点  $v$ , 否则随机选择集合  $V^1$  中一个节点  $v$ , 转到步骤S607;

S607, 如果  $d_v \geq 2 \vee v \notin \bar{V}_1$ , 则随机选择一条与节点  $v$  相邻的边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E')$ ,  $E' \leftarrow E' / \{e\}$ , 如果  $d_v = 0$ , 则更新  $V_1^{L'} \leftarrow V_1^{L'} / \{v\}$ , 如果  $d_v < 2 \wedge v \in \bar{V}_1$ , 则此步不作任何操作; 转到步骤S609;

S608, 从集合  $\{(w, v) \in E' \mid w \in V'_2 \wedge v \in V_1^{H'}\} / \bar{V}_1$  中随机选择一条边  $e$ , 并更新二分  $G' = (V'_1, V'_2, E') : E' \leftarrow E' / \{e\}$ ; 如果  $d_v = 0$ , 则更新  $V_1^{H'} \leftarrow V_1^{H'} / \{v\}$ ; 转到步骤S609;

S609, 更新  $t \leftarrow t + 1$ , 转到步骤S604;

S610, 输出: 采样子图  $G' = (V'_1, V'_2, E')$ ; 其中, 输入的节点集  $\bar{V}_1$  被包含于采样子图  $G' = (V'_1, V'_2, E')$ , 即  $\bar{V}_1 \subseteq V'_1$ ; 图  $G$  中节点集  $V_2$  包含节点的度均为1, 图  $G'$  中节点集  $V'_2$  包含节点的度均为1;

二分图采样4单元,用于调用二分图采样方法4对二分图 $Q_3$ -PI进行采样,抽取部分节点和边构成采样子图;

其中,所述调用二分图采样方法4包括如下步骤:

S701,输入:二分图 $G=(V_1, V_2, E)$ ,其中 $V_1$ 和 $V_2$ 为两个节点集、 $E$ 为边集,多对一映射函数 $\psi:V_1 \rightarrow W$ ,设 $\psi(v_1)$ 为节点集 $W$ 中与节点 $v_1 \in V_1$ 对应的唯一节点,对任意节点 $w \in W$ 预选设定参数 $L(w)$ ,预选给定节点度分布 $\{g(d) \mid d=1, 2, \dots\}$ ,预先设定参数 $T, R$ ,规模缩减率 $R_R$ ;转到步骤S702;

S702,将 $V_2$ 中全部节点按度从大到小顺利排列,并抽取前 $n_h$ 个最大度节点构成节点子集 $V_2^{LH}$ ,其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ;转到步骤S703;

S703,计算 $V_2^{LL} = V_2 / V_2^{LH}$ ,并将二分图 $G$ 分解为 $G^{LH}=(V_1, V_2^{LH}, E^{LH})$ 和 $G^{LL}=(V_1, V_2^{LL}, E^{LL})$ ,其中 $E^{LH}=\{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LH}\}$ 且 $E^{LL}=\{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ;转到步骤S704;

S704,设 $d_1, d_2, \dots, d_s$ 为二分图 $G$ 中节点集 $V_1$ 的全部节点度,并设 $d_{\max}$ 为节点集 $V_2^{LL}$ 中的最大节点度,初始化二分图 $G'=(V'_1, V'_2, E') \leftarrow G=(V_1, V_2, E)$ ,初始化二分图 $G^{LH'}=(V'_1, V_2^{LH'}, E^{LH'}) \leftarrow G^{LH}=(V_1, V_2^{LH}, E^{LH})$ ,初始化二分图 $G^{LL'}=(V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL}=(V_1, V_2^{LL}, E^{LL})$ ,计算二分图 $G^{LL'}$ 中节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) \mid d=1, 2, \dots, d_{\max}\}$ ,其中, $g'(d)$ 表示 $V_2^{LL'}$ 中度为 $d$ 的节点比率,计算最终需要删除边的总数 $E_d = ||E|| \cdot R_R$ ,其中,初始化变量 $t=0, \eta=0$ ;转到步骤S705;

S705,如果 $t < E_d$ ,转到步骤S706;否则二分图 $G'=(V'_1, V'_2, E')$ ,其中 $V'_2 = V_2^{LH'} \cup V_2^{LL'}$ 且 $E' = E^{LH'} \cup E^{LL'}$ ,为最终输出结果,转到步骤S713;

S706,如果 $\eta > 0 \wedge ||\tilde{V}_1|| > 0$ ,则更新 $\eta \leftarrow \eta - 1$ ,计算两个可选边集 $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\}$ 和 $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ ,

$$\text{其中 } \tilde{V}_1 = \left\{ v_1 \in V_1' \mid \frac{||N'(\psi(v_1))||}{L(\psi(v_1))} > 1 \wedge d_v = 1 \right\}, \text{ 且集合 } N'(\psi(v_1)) = \{v \mid v \in V_1' \wedge \psi(v) = \psi(v_1)\},$$

$= \psi(v_1)\}$ ,

并转到步骤S709,否则,转到步骤S707;

S707,如果 $||S = \{v \mid v \in V_1' \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}|| > 0$ ,选择集合 $S$ 中一个节点 $v$ ,设该节点的度为 $d = d_v$ ;否则以式 $p_k = \sum_{i=1}^k f(d_i)$ 定义的离散概率分布 $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k=1, 2, \dots, s$ )随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ ,其中 $d_1 > d_2 > \dots > d_s > 0$ , $f(d_1), f(d_2), \dots, f(d_s)$ 表示节点集 $V_1$ 中度为 $d_k$  ( $1 \leq k \leq s$ )的节点比率,转到步骤S708;

S708,如果 $d = 1 \wedge ||\tilde{V}_1|| = 0$ ,则更新 $\eta \leftarrow \eta + 1$ ,转到步骤S712;如果 $d = 1 \wedge ||\tilde{V}_1|| > 0$ ,则计算两个可选边集 $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\}$ 和 $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ ,转到步骤S709;如果 $d > 1$ ,则计算两个可选边集

$E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V_1' \wedge v_2 \in V_2^{LH'} \wedge d_v = d\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V_1' \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}$ , 转到步骤S709;

S709, 如果  $\|E_o^{LH}\| = 0$ , 转到步骤S710; 如果  $\|E_o^{LL}\| = 0$ , 转到步骤S711; 如果  $\|E_o^{LH}\| > 0 \wedge \|E_o^{LL}\| > 0$ , 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤S710, 以概率  $1 - P^{LL'}$  转到步骤S711,

其中,  $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ ,  $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ ,

$$m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R), m^{LH} = E_d \cdot m^{LL};$$

S710, 假设  $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ , 计算  $D^h = \{d \in D \mid g'(d) - g(d) > 1 / \|V_2^{LL'}\|\}$ , 其中D为集合  $V_2^{LL'}$  中节点  $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合; 如果  $||D^h|| = 0$ , 则更新  $D^h \leftarrow \{\arg \max_{d \in D} \{g'(d) - g(d)\}\}$ ;

首先, 计算集合  $E_o^{LL}$  的一个子集:  $E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} \in D^h\}$ ,

假设  $E_a = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_q^l, v_q^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_q^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_q^h \in V_2^{LL'}$ ; 计算  $\psi' = \arg \max_{\psi(v)} \{||N'(\psi(v))|| / L(\psi(v))\}$ , 其中  $V = \{v_1^l, v_2^l, \dots, v_q^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$ ,

然后, 计算集合  $E_a$  的一个子集,  $E_b = \{(v_i^l, v_i^h) \in E_a \mid \psi(v_i^h) = \psi'\}$ ,

最后, 从集合  $E_b$  中随机选择一条边  $e$ , 并将该边从二分图  $G^{LL'} = (V_1', V_2^{LL'}, E^{LL'})$  和  $G' = (V_1', V_2', E')$  中删除, 即更新  $E^{LL'} \leftarrow E^{LL'} / \{e\}$  且  $E' \leftarrow E' / \{e\}$ , 更新二分  $G^{LL'}$  中  $V_2^{LL'}$  的节点度分布  $\{g'(d) \mid d = 1, 2, \dots, d_{\max}\}$ , 其中  $g'(d)$  表示  $V_2^{LL'}$  中度为  $d$  的节点比率, 转到步骤S712;

S711, 假设  $E_o^{LH} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LH'}$ ; 计算  $\psi' = \arg \max_{\psi(v)} \{||N'(\psi(v))|| / L(\psi(v))\}$ , 其中  $V = \{v_1^l, v_2^l, \dots, v_r^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$ ,

计算集合  $E_o^{LH}$  的一个子集:  $E_b = \{(v_i^l, v_i^h) \in E_o^{LH} \mid \psi(v_i^h) = \psi'\}$ ,

并从集合  $E_b$  中随机选择一条边  $e$ , 并将该边从二分图  $G^{LH'} = (V_1', V_2^{LH'}, E^{LH'})$  和  $G' = (V_1', V_2', E')$  中删除, 即更新  $E^{LH'} \leftarrow E^{LH'} / \{e\}$  且  $E' \leftarrow E' / \{e\}$ , 转到步骤S712;

S712, 更新  $t \leftarrow t + 1$ , 转到步骤S705;

S713, 输出: 采样子图  $G' = (V_1', V_2', E')$ ; 其中, 对  $\forall w \in W$ , 集合  $N'(w) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = w\}$ , 其中  $W$  为步骤S701的输入,  $V_1' \subseteq V_1$ , S701至S713期望对  $\forall w \in W$  满足  $||N'(w)||$  接近于  $L(w)$ ;

子图合并单元,用于对得到的所有采样子图进行合并,得到互联网外围拓扑所有采样子图的合并图;

内核图采样单元,用于调用内核图采样方法对内核图进行采样,从而抽取部分节点和边构成采样子图;

其中,所述调用内核图采样方法包括如下步骤:

S901,输入原始互联网拓扑 $G=(V,E)$ 、合并子图 $G'_m$ 和内核子图 $G'_{Core}=(V'_{Core},E'_{Core})$ ;

S902,计算原始拓扑 $G$ 的平均节点度 $d_{avg}$ ,计算合并子图 $G'_m$ 的边数 $m(G'_m)$ 和节点数 $n(G'_m)$ ,将所有边 $(w,v) \in E'_{Core}$ 按照 $\max(d_w(G'_{Core}),d_v(G'_{Core}))$ 不降顺序排列;

S903,从边集 $E'_{Core}$ 中删除排在前面的 $||E'_{Core}||-d_{avg} \cdot n(G'_m)/2+m(G'_m)$ 条边,并将剩余边组成的集合定义为 $E''_{Core}$ ;

S904, $E''_{Core}$ 为最终输出的内核边集;

S905,输出内核边集 $E''_{Core}$ ;

采样子图获取单元,用于对得到的外围拓扑采样子图的合并图和内核拓扑的采样子图进行合并,从而得到真实世界互联网拓扑图的采样子图;

仿真测试单元,依据得到的真实世界互联网拓扑图的规模缩减采样子图,配置互联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境。

## 一种互联网测试床拓扑结构大比例规模缩减方法及装置

### 技术领域

[0001] 本公开涉及数据挖掘应用技术领域,具体涉及一种互联网测试床拓扑结构大比例规模缩减方法及装置。

### 背景技术

[0002] 路由协议、资源定位等互联网技术的开发、测试与评估,需要构建相应的仿真测试环境。面对真实互联网网络中每年快速增长的庞大节点规模,相应测试床的构建必须采用规模缩减的方式,以实现构建成本的缩减以及仿真资源利用效率和测试时间运行效率的提高。虽然已有图采样技术已被应用于自治系统级互联网拓扑的规模缩减,但是这些技术的应用范围广泛,没有考虑互联网拓扑独有的结构特征,难以在保持重要拓扑属性稳定性的条件下达到70%以上的规模缩减率(定义为缩减过程删除的节点数占原始拓扑中节点数的比率)。通用性越好的工具,难以适用于专业性较强的领域。因此,本发明关注于自治系统级互联网拓扑独有的二分图结构特征,采用二分图的分解、采样与合并策略,能够在保持重要拓扑属性稳定性的条件下实现真实世界自治系统级互联网拓扑96%以上的规模缩减率。依据压缩规模后的拓扑结构,在小规模互联网测试床上配置拓扑连接关系,可以实现真实大规模拓扑在小规模测试床的等效部署。

### 发明内容

[0003] 本公开的目的是针对现有技术的不足,本公开提出一种互联网测试床拓扑结构大比例规模缩减方法及装置,所述方法包括如下步骤:

[0004] S1,获取真实世界互联网拓扑结构的探测数据,以简单无向图 $G=(V,E)$ 建模自治系统级互联网拓扑结构,其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域,集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径;

[0005] S2,采用互联网拓扑结构模型分解互联网拓扑图 $G=(V,E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $Q_{I_1-RI}$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ ,一个匹配图 $BI-BI$ 和一个内核图 $Core$ ;

[0006] S3,调用二分图采样方法1对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $Q_{I_1-RI}$ 进行采样,从而抽取部分节点和边构成采样子图;

[0007] S4,调用匹配图采样方法对匹配图 $BI-BI$ 进行采样,抽取部分节点和边构成采样子图;

[0008] S5,调用二分图采样方法2对二分图 $Q_4-RI_1$ 进行采样,抽取部分节点和边构成采样子图;

[0009] S6,调用二分图采样方法3对二分图 $QI-PI$ 和二分图 $Q-P$ 进行采样,抽取部分节点和边构成采样子图;

[0010] S7,调用二分图采样方法4对二分图 $Q_3-PI$ 进行采样,抽取部分节点和边构成采样子图;

[0011] S8,对S3至S7得到的所有采样子图进行合并,得到这些采样子图的合并图;

[0012] S9,调用内核图采样方法对内核图Core进行采样,从而抽取部分节点和边构成采样子图;

[0013] S10,对S8得到的合并图和S9得到的采样子图进行合并,从而得到真实世界互联网拓扑图 $G=(V,E)$ 的采样子图;

[0014] S11,依据S10得到的真实世界互联网拓扑图的规模缩减采样子图,配置互联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境;

[0015] 进一步地,对图进行采样,是指从该图中抽取部分节点和边,从而生成由这些被抽取节点和边构成的一个子图。

[0016] 其中,所述互联网拓扑结构模型分解为八类节点集,其中节点集OI表示系统噪声:

$$[0017] \begin{cases} P = \{v \in V \mid d_v = 1\} \\ Q = \{v \in V \mid \exists w, (v, w) \in E, w \in P\} \\ PI = \{v \in V_I \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) > 1\} \\ QI = \{v \in V_I \mid \exists w, (v, w) \in E_I, w \in PI\} \\ RI = \{v \in V_I \mid d_I(v) \geq 2 \wedge \forall (v, w) \in E_I, w \in QI\} \\ BI = \{v \in V_I \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) = 1\} \\ II = \{v \in V_I \mid d_I(v) = 0\} \\ OI = V_I / (PI \cup QI \cup RI \cup BI \cup II) \end{cases}$$

[0018] 其中,其中V和E分别为互联网拓扑图 $G=(V,E)$ 的节点集和边集,v为拓扑图 $G=(V,E)$ 中的节点, $d_v$ 为拓扑图 $G=(V,E)$ 中节点v的度; $V_I, E_I$ :分别为子图 $G_I=(V_I, E_I)$ 的节点集和边集; $G_I=(V_I, E_I)$ 为拓扑图 $G=(V,E)$ 中,由节点集 $R(G)=V/(P \cup Q)$ 和两个端点均属于 $R(G)$ 的所有边构成边集生成的子图; $d_I(v)$ 为子图 $G_I=(V_I, E_I)$ 中节点v的度。

[0019] 进一步地,互联网拓扑结构模型的八类节点集节点之间的连接关系为:每个P节点仅且必被连接至唯一的Q节点,每个II节点仅且必被连接至至少两个Q节点,每个PI节点仅且必被连接至唯一的QI节点和至少一个Q节点,每个BI节点仅且必被连接至另外一个BI节点和至少一个Q节点,每个RI节点仅且必被连接至至少两个QI节点,此外QI和RI节点可能被连接至Q节点,由连接关系将自治系统级互联网拓扑分解为七个二分图 $Q_1-II, Q-P, Q_2-BI, Q_3-PI, QI-PI, QI_1-RI$ 和 $Q_4-RI_1$ ,其中节点集 $Q_1 = \{v \in Q \mid \exists w \in II, (v, w) \in E\}$ 、 $Q_2 = \{v \in Q \mid \exists w \in BI, (v, w) \in E\}$ 、 $Q_3 = \{v \in Q \mid \exists w \in PI, (v, w) \in E\}$ 、 $Q_4 = \{v \in Q \mid \exists w \in RI, (v, w) \in E\}$ 、

$QI_1 = \{v \in QI \mid \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI \mid \exists v \in Q, (v, w) \in E\}$ ,一个匹配图BI-BI和一个内核图Core。

[0020] 进一步地,二分图是指图中全部节点可以被划分至两个不同的节点集,且图中每条边的两个端点必定分别属于这两个不同的节点集。二分图 $Q_1-II$ 是指由节点集 $Q_1$ 、节点集II和两个端点分别属于 $Q_1$ 和II的所有边构成边集生成的二分图结构。二分图 $Q-P, Q_2-BI, Q_3-PI, QI-PI, QI_1-RI$ 和 $Q_4-RI_1$ 的物理意义解释方法与二分图 $Q_1-II$ 相同。匹配图BI-BI是由节点集BI和两个端点都属于BI的所有边构成边集生成的图结构。内核图Core是由节点集 $Q \cup QI$ 和连接该节点集内部节点的所有边构成边集生成的图结构。

[0021] 进一步地,在S2中,互联网拓扑图分解方法包括如下步骤:

[0022] S201, 读取互联网拓扑图  $G = (V, E)$  的七个节点集  $P, Q, II, BI, QI, PI$  和  $RI$ ;

[0023] S202, 计算二分图  $Q_1-II: G_{Q_1-II} = (II, Q_1, E_{Q_1-II})$ , 其中  $Q_1 = \{v \in Q \mid \exists w \in II, (v, w) \in E\}$  且  $E_{Q_1-II} = \{e = (v, w) \mid e \in E \wedge v \in II \wedge w \in Q_1\}$ ;

[0024] S203, 计算二分图  $Q_2-BI: G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$ , 其中  $Q_2 = \{v \in Q \mid \exists w \in BI, (v, w) \in E\}$  且  $E_{Q_2-BI} = \{e = (v, w) \mid e \in E \wedge v \in BI \wedge w \in Q_2\}$ ;

[0025] S204, 计算二分图  $QI_1-RI: G_{QI_1-RI} = (RI, QI_1, E_{QI_1-RI})$ , 其中  $QI_1 = \{v \in QI \mid \exists w \in RI, (v, w) \in E\}$  且  $E_{QI_1-RI} = \{e = (v, w) \mid e \in E \wedge v \in RI \wedge w \in QI_1\}$ ;

[0026] S205, 计算二分图  $Q_4-RI_1: G_{Q_4-RI_1} = (RI_1, Q_4, E_{Q_4-RI_1})$ , 其中  $Q_4 = \{v \in Q \mid \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI \mid \exists v \in Q, (v, w) \in E\}$  且  $E_{Q_4-RI_1} = \{e = (v, w) \mid e \in E \wedge v \in RI_1 \wedge w \in Q_4\}$ ;

[0027] S206, 计算二分图  $QI-PI: G_{QI-PI} = (QI, PI, E_{QI-PI})$ , 其中  $E_{QI-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in QI\}$ ;

[0028] S207, 计算二分图  $Q_3-PI: G_{Q_3-PI} = (PI, Q_3, E_{Q_3-PI})$ , 其中  $Q_3 = \{v \in Q \mid \exists w \in PI, (v, w) \in E\}$  且  $E_{Q_3-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in Q_3\}$ ;

[0029] S208, 计算二分图  $Q-P: G_{Q-P} = (Q, P, E_{Q-P})$ , 其中  $E_{Q-P} = \{e = (v, w) \mid e \in E \wedge v \in P \wedge w \in Q\}$ ;

[0030] S209, 计算匹配图  $BI-BI: G_{BI-BI} = (BI, E_{BI-BI})$ , 其中  $E_{BI-BI} = \{e = (v, w) \mid e \in E \wedge v, w \in BI\}$ ;

[0031] S210, 计算内核图  $Core: G_{Core} = (V_{Core}, E_{Core})$ , 其中  $V_{Core} = Q \cup QI$  且  $E_{Core} = \{e = (v, w) \mid e \in E \wedge v, w \in V_{Core}\}$ ;

[0032] S211, 输出七个二分图  $G_{Q_1-II}$ 、 $G_{Q_2-BI}$ 、 $G_{QI_1-RI}$ 、 $G_{Q_4-RI_1}$ 、 $G_{QI-PI}$ 、 $G_{Q_3-PI}$ 、 $G_{Q-P}$ , 一个匹配图  $G_{BI-BI}$  和一个内核图  $G_{Core}$ 。

[0033] 其中,  $\in$  表示“属于”;  $\exists$  表示“存在”;  $\wedge$  表示“逻辑与”;  $\cup$  表示集合的并集;  $V$  和  $E$  分别表示互联网拓扑图  $G = (V, E)$  的节点集和边集;  $P, Q, II, BI, QI, PI$  和  $RI$  分别为互联网拓扑结构模型分解得到的七类节点集。

[0034] 进一步地, 在S3中, 所述调用二分图采样方法1包括如下步骤:

[0035] S301, 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1$  和  $V_2$  为两个节点集、 $E$  为边集, 参数  $T, R$ , 规模缩减率  $R_R$ ; 转到步骤S302;

[0036] S302, 将  $V_2$  中全部节点按度从大到小顺利排列, 并抽取前  $n_h$  个最大度节点构成节点子集  $V_2^{LH}$ , 其中这  $n_h$  个节点中相同度重复数的最大值为  $T$ ; 转到步骤S303;

[0037] 进一步地,  $n_h$  的取值由节点集  $V_2$  中节点的度和参数  $T$  唯一确定。

[0038] S303, 计算  $V_2^{LL} = V_2 / V_2^{LH}$ , 并将二分图  $G$  分解为  $G^{LH} = (V_1, V_2^{LH}, E^{LH})$  和  $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 其中  $E^{LH} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LH}\}$  且  $E^{LL} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ; 转到步骤S304;

[0039] S304, 设  $d_1, d_2, \dots, d_s$  为图  $G$  中节点集  $V_1$  的全部节点度, 并设  $d_{max}$  为节点集  $V_2^{LL}$  中的最

大节点度,计算节点集 $V_2^{LL}$ 的度分布 $\{g(d) \mid d=1,2,\dots,d_{\max}\}$ ,其中 $g(d)$ 表示 $V_2^{LL}$ 中度为 $d$ 的节点比率,初始化二分图 $G'=(V'_1, V'_2, E') \leftarrow G=(V_1, V_2, E)$ ,初始化二分图 $G^{LH'}=(V'_1, V_2^{LH'}, E^{LH'}) \leftarrow G^{LH}=(V_1, V_2^{LH}, E^{LH})$ ,初始化二分图

$G^{LL'}=(V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL}=(V_1, V_2^{LL}, E^{LL})$ ,初始化二分图 $G^{LL'}$ 中节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) \mid d=1,2,\dots,d_{\max}\} \leftarrow \{g(d) \mid d=1,2,\dots,d_{\max}\}$ ,计算需删除边的总数 $E_d=||E|| \cdot R_R$ ,初始化变量 $t=0$ ,转到步骤S305;

[0040] S305,如果 $t < E_d$ ,转到步骤S306;否则 $G'=(V'_1, V'_2, E')$ ,其中 $V'_2=V_2^{LH'} \cup V_2^{LL'}$ 且 $E'=E^{LH'} \cup E^{LL'}$ ,为最终输出结果,转到步骤S311;

[0041] S306,如果 $||S=\{v \mid v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}|| > 0$ ,任意选择集合 $S$ 中一个节点 $v$ ,并计算两个可选边集 $E_o^{LH}=\{(v, v_2) \in E^{LH'} \mid v_2 \in V_2^{LH'}\}$ 和 $E_o^{LL}=\{(v, v_2) \in E^{LL'} \mid v_2 \in V_2^{LL'}\}$ ,如果 $||E_o^{LH}||=0$ ,转到步骤S308;如果 $||E_o^{LL}||=0$ ,转到步骤S309;否则 $||E_o^{LH}|| > 0$ 且 $||E_o^{LL}|| > 0$ ,以概率 $P^{LL'}=\gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$ 转到步骤S308,并以概率 $1-P^{LL'}$ 转到步骤S309,其中, $\gamma^{LL'}=m^{LL}-(||E^{LL}||-||E^{LL'}||)$ , $\gamma^{LH'}=m^{LH}-(||E^{LH}||-||E^{LH'}||)$ , $m^{LL}=\frac{E_d \cdot ||E^{LL}||}{||E^{LH} \cup E^{LL}||} \cdot (1+R)$ , $m^{LH}=E_d-m^{LL}$ ;转到步骤

S307;

[0042] S307,如果 $S=\Phi$ ,以式 $p_k=\sum_{i=1}^k f(d_i)$ 定义的离散概率分布 $p_k / \sum_{i=1,2,\dots,s} p_i$  ( $k=1,2,\dots,s$ )随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ ,其中 $d_1 > d_2 > \dots > d_s > 0$ , $f(d_1), f(d_2), \dots, f(d_s)$ 表示节点集 $V_1$ 中度为 $d_k$  ( $1 \leq k \leq s$ )的节点比率;计算两个可选边集 $E_o^{LH}=\{(v, v_2) \in E^{LH'} \mid v \in V'_1 \wedge v_2 \in V_2^{LH'} \wedge d_v=d\}$ , $E_o^{LL}=\{(v, v_2) \in E^{LL'} \mid v \in V'_1 \wedge v_2 \in V_2^{LL'} \wedge d_v=d\}$ ,如果 $||E_o^{LH}||=0$ ,转到步骤S308;如果 $||E_o^{LL}||=0$ ,转到步骤S309,否则,以概率 $P^{LL'}=\gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$ 转到步骤S308,并以概率 $1-P^{LL'}$ 转到步骤S309;

[0043] S308,假设 $E_o^{LL}=\{(v'_1, v_1^h), (v'_2, v_2^h), \dots, (v'_r, v_r^h)\}$ ,其中 $v'_1, v'_2, \dots, v'_r \in V'_1$ 且 $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ ,计算 $d^h=\arg \max_{d \in D} \{g'(d)-g(d)\}$ ,其中 $D$ 为集合 $V_2^{LL'}$ 中节点 $v_1^h, v_2^h, \dots, v_r^h$ 的度组成的集合,并计算 $E_o^{LL}$ 的子集 $E_a=\{(v'_i, v_i^h) \in E_o^{LL} \mid d_{v_i^h}=d^h\}$ ,从 $E_a$ 中均等概率分布地采样一条边 $e$ ,并将该边从图 $G^{LL'}=(V'_1, V_2^{LL'}, E^{LL'})$ 和 $G'=(V'_1, V'_2, E')$ 中删除,即更新 $E^{LL'} \leftarrow E^{LL'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,更新 $G^{LL'}$ 中 $V_2^{LL'}$ 的节点度分布 $\{g'(d) \mid d=1,2,\dots,d_{\max}\}$ ,其中 $g'(d)$ 表示 $V_2^{LL'}$ 中度为 $d$ 的节点比率,转到步骤S310;

[0044] S309,从 $E_o^{LH}$ 中均等概率分布地采样一条边 $e$ ,并将该边从二分图 $G^{LH'}=(V'_1, V_2^{LH'}, E^{LH'})$ 和 $G'=(V'_1, V'_2, E')$ 中删除,即更新 $E^{LH'} \leftarrow E^{LH'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,转到步骤S310;

[0045] S310,更新 $t \leftarrow t+1$ ,转到步骤S305;

[0046] S311,输出:采样子图 $G' = (V'_1, V'_2, E')$ 。

[0047] 其中,  $||S||$  表示集合S的势,即集合S包含元素的总数;  $x \leftarrow y$  表示将右边的函数或数值y赋值给左边的变量x;  $d_v$  表示节点v在图 $G' = (V'_1, V'_2, E')$ 中的度;  $A/B$ : 如果A和B为数值,则表示除法,如果A和B为集合,则表示从集合A中删除集合B的所有元素;  $d^h = \arg \max_{d \in D} \{g'(d) - g(d)\}$  表示将 $g'(d) - g(d)$ 取最大值时参数d的取值赋值给 $d^h$ ;  $\in$  表示“属于”;  $\notin$  表示“不属于”;  $\wedge$  表示“逻辑与”; 节点的度是指与该节点相邻边的总数;  $\cup$  表示集合的并集; 输入参数规模缩减率 $R_R$ 被定义为需要从输入的二分图中删除边的总数与输入的二分图包含边数的比率;  $\Phi$  表示空集,即包含元素个数为零的集合。

[0048] 进一步地,在S6中,所述调用二分图采样方法3包括如下步骤:

[0049] S601,输入:二分图 $G = (V_1, V_2, E)$ ,其中 $V_1$ 和 $V_2$ 为两个节点集、E为边集、节点集 $V_2$ 包含节点的度均为1,节点集 $\bar{V}_1 \subseteq V_1$ ,参数T,规模缩减率 $R_R$ ;转到步骤S602;

[0050] S602,将 $V_1$ 中全部节点按度从大到小顺利排列,并抽取前 $n_h$ 个最大度节点构成节点子集 $V_1^H$ ,其中这 $n_h$ 个节点中相同度重复数的最大值为T;转到步骤S603;

[0051] S603,计算 $V_1^L = V_1 / V_1^H$ ,设 $d_1, d_2, \dots, d_s$ 为二分图G中节点集 $V_1$ 的全部节点度,初始化二分图 $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ ,初始化节点集 $V_1^{L'} \leftarrow V_1^L$ ,初始化节点集 $V_1^{H'} \leftarrow V_1^H$ ,计算最终需要删除边的总数 $E_d = ||E|| R_R$ ,初始化变量 $t = 0$ ;转到步骤S604;

[0052] S604,如果 $t < E_d$ ,则以概率 $P = \gamma^{L'} / (\gamma^{H'} + \gamma^{L'})$ 转到步骤S605,并以概率 $1-P$ 转到步骤S608;否则, $G' = (V'_1, V'_2, E')$ ,其中 $V'_1 = V_1^{H'} \cup V_1^{L'}$ ,为最终输出结果,转到步骤S610;其中,

$$[0053] \quad \gamma^{L'} = m^L - \left( \sum_{v \in V_1^L} d_v - \sum_{v \in V_1^{L'}} d_v \right), \quad \gamma^{H'} = m^H - \left( \sum_{v \in V_1^H} d_v - \sum_{v \in V_1^{H'}} d_v \right),$$

$$[0054] \quad m^L = \frac{E_d \cdot \sum_{v \in V_1^L} d_v}{\sum_{v \in V_1^H \cup V_1^L} d_v} \text{ 且 } m^H = E_d - m^L;$$

[0055] S605,如果 $\left\| S = \left\{ v \mid v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\} \right\} / \tilde{V}_1 \right\| > 0$ ,其中 $\tilde{V}_1 = \{v \mid d_v = 1 \wedge v \in \bar{V}_1\}$ ,则

随机选择集合S中一个节点v,转到步骤S607,否则,转到步骤S606;

[0056] S606,以式 $p_k = \sum_{i=1}^k f(d_i)$ 定义的离散概率分布 $p_k / \sum_{i=1,2,\dots,s} p_i$  ( $k=1,2,\dots,s$ ),随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ ,其中, $d_1 > d_2 > \dots > d_s > 0$ , $f(d_1), f(d_2), \dots, f(d_s)$ 表示节点集 $V_1$ 在二分图 $G = (V_1, V_2, E)$ 中度为 $d_k$  ( $1 \leq k \leq s$ )的节点比率,计算两个节点集 $V^1 = \{v \mid v \in V_1^{L'} \cap \bar{V}_1 \wedge d_v = d\}$ 和 $V^2 = \{v \mid v \in V_1^{L'} / \bar{V}_1 \wedge d_v = d\}$ ,如果 $||V^2|| > 0$ ,则随机选择集合 $V^2$ 中一个节点v,否则随机选择集合 $V^1$ 中一个节点v,转到步骤S607;

[0057] S607,如果 $d_v \geq 2 \vee v \notin \bar{V}_1$ ,则随机选择一条与节点v相邻的边e,并更新二分图 $G' = (V'_1, V'_2, E')$ , $E' \leftarrow E' / \{e\}$ ,如果 $d_v = 0$ ,则更新 $V_1^{L'} \leftarrow V_1^{L'} / \{v\}$ ,如果 $d_v < 2 \wedge v \in \bar{V}_1$ ,则此步不作任何操作;转到步骤S609;

[0058] S608, 从集合  $\{(w,v) \in E' \mid w \in V'_2 \wedge v \in V_1^{H'}\} / \tilde{V}_1$  中随机选择一条边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E') : E' \leftarrow E' / \{e\}$ ; 如果  $d_v = 0$ , 则更新  $V_1^{H'} \leftarrow V_1^{H'} / \{v\}$ ; 转到步骤S609;

[0059] S609, 更新  $t \leftarrow t+1$ , 转到步骤S604;

[0060] S610, 输出: 采样子图  $G' = (V'_1, V'_2, E')$ 。

[0061] 进一步地, 输入的节点集  $\tilde{V}_1$  被包含于采样子图  $G' = (V'_1, V'_2, E')$ , 即  $\tilde{V}_1 \subseteq V'_1$ ; 图  $G$  中节点集  $V_2$  包含节点的度均为1, 图  $G'$  中节点集  $V'_2$  包含节点的度均为1;

[0062] 其中,  $||S||$  表示集合  $S$  的势, 即集合  $S$  包含元素的总数;  $x \leftarrow y$  表示将右边的函数或数值  $y$  赋值给左边的变量  $x$ ; 当节点  $v \in V_1$  或  $v \in V_1^H$  或  $v \in V_1^L$ ,  $d_v$  表示节点  $v$  在图  $G = (V_1, V_2, E)$  的度, 其他情况或没有特定约束下,  $d_v$  表示节点  $v$  在图  $G' = (V'_1, V'_2, E')$  的度;  $A/B$ : 如果  $A$  和  $B$  为数值, 则表示除法, 如果  $A$  和  $B$  为集合, 则表示从集合  $A$  中删除集合  $B$  的所有元素;  $\in$  表示“属于”;  $\notin$  表示“不属于”;  $\wedge$  表示“逻辑与”;  $\vee$  表示“逻辑或”; 节点的度是指与该节点相邻边的总数;  $A \subseteq B$  表示集合  $A$  被包含于集合  $B$ ;  $\cup$  表示集合的并集;  $\cap$  表示集合的交集; 输入参数规模缩减率  $R_R$  被定义为需要从输入的二分图中删除边的总数与输入的二分图包含边数的比率。

[0063] 进一步地, 在S7中, 所述调用二分图采样方法4包括如下步骤:

[0064] S701, 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1$  和  $V_2$  为两个节点集、 $E$  为边集, 多对一映射函数  $\psi: V_1 \rightarrow W$ , 设  $\psi(v_1)$  为节点集  $W$  中与节点  $v_1 \in V_1$  对应的唯一节点, 对任意节点  $w \in W$  预选设定参数  $L(w)$ , 预选给定节点度分布  $\{g(d) \mid d=1, 2, \dots\}$ , 预先设定参数  $T, R$ , 规模缩减率  $R_R$ ; 转到步骤S702;

[0065] S702, 将  $V_2$  中全部节点按度从大到小顺利排列, 并抽取前  $n_h$  个最大度节点构成节点子集  $V_2^{HH}$ , 其中这  $n_h$  个节点中相同度重复数的最大值为  $T$ ; 转到步骤S703;

[0066] S703, 计算  $V_2^{LL} = V_2 / V_2^{HH}$ , 并将二分图  $G$  分解为  $G^{HH} = (V_1, V_2^{HH}, E^{HH})$  和  $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 其中  $E^{HH} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{HH}\}$  且  $E^{LL} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ ; 转到步骤S704;

[0067] S704, 设  $d_1, d_2, \dots, d_s$  为二分图  $G$  中节点集  $V_1$  的全部节点度, 并设  $d_{\max}$  为节点集  $V_2^{LL}$  中的最大节点度, 初始化二分图  $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ , 初始化二分图  $G^{HH'} = (V'_1, V_2^{HH'}, E^{HH'}) \leftarrow G^{HH} = (V_1, V_2^{HH}, E^{HH})$ , 初始化二分图

$G^{LL'} = (V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL} = (V_1, V_2^{LL}, E^{LL})$ , 计算二分图  $G^{LL'}$  中节点集  $V_2^{LL'}$  的度分布  $\{g'(d) \mid d=1, 2, \dots, d_{\max}\}$ , 其中,  $g'(d)$  表示  $V_2^{LL'}$  中节点度为  $d$  的节点比率, 计算最终需要删除边的总数  $E_d = ||E|| R_R$ , 其中, 初始化变量  $t=0, \eta=0$ ; 转到步骤S705;

[0068] S705, 如果  $t < E_d$ , 转到步骤S706; 否则二分图  $G' = (V'_1, V'_2, E')$ , 其中  $V'_2 = V_2^{HH'} \cup V_2^{LL'}$  且  $E' = E^{HH'} \cup E^{LL'}$ , 为最终输出结果, 转到步骤S713;

[0069] S706, 如果  $\eta > 0 \wedge ||\tilde{V}_1|| > 0$ , 则更新  $\eta \leftarrow \eta-1$ , 计算两个可选边集  $E_o^{HH} = \{(v, v_2) \in E^{HH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{HH'}\}$

和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ ,

[0070] 其中  $\tilde{V}_1 = \left\{ v_1 \in V_1' \mid \frac{\|N'(\psi(v_1))\|}{L(\psi(v_1))} > 1 \wedge d_v = 1 \right\}$ , 且集合  $N'(\psi(v_1)) = \{v \mid v \in V_1' \wedge \psi(v) = \psi(v_1)\}$ ,

[0071] 并转到步骤S709, 否则, 转到步骤S707;

[0072] S707, 如果  $\|S = \{v \mid v \in V_1' \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} \| > 0$ , 选择集合S中一个节点v, 设该节点的度为  $d = d_v$ ; 否则以式  $p_k = \sum_{i=1}^k f(d_i)$  定义的离散概率分布  $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k = 1, 2, \dots, s$ ) 随机采样一个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  中度为  $d_k$  ( $1 \leq k \leq s$ ) 的节点比率, 转到步骤S708;

[0073] S708, 如果  $d = 1 \wedge \|\tilde{V}_1\| = 0$ , 则更新  $\eta \leftarrow \eta + 1$ , 转到步骤S712; 如果  $d = 1 \wedge \|\tilde{V}_1\| > 0$ , 则计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ , 转到步骤S709; 如果  $d > 1$ , 则计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V_1' \wedge v_2 \in V_2^{LH'} \wedge d_v = d\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V_1' \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}$ , 转到步骤S709;

[0074] S709, 如果  $\|E_o^{LH}\| = 0$ , 转到步骤S710; 如果  $\|E_o^{LL}\| = 0$ , 转到步骤S711; 如果  $\|E_o^{LH}\| > 0 \wedge \|E_o^{LL}\| > 0$ , 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转到步骤S710, 以概率  $1 - P^{LL'}$  转到步骤S711,

[0075] 其中,  $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ ,  $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ ,

[0076]  $m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R)$ ,  $m^{LH} = E_d - m^{LL}$ ;

[0077] S710, 假设  $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ , 计算  $D^h = \{d \in D \mid g'(d) - g(d) > 1 / \|V_2^{LL'}\|\}$ , 其中D为集合  $V_2^{LL'}$  中节点  $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合; 如果  $||D^h|| = 0$ , 则更新  $D^h \leftarrow \{\arg \max_{d \in D} \{g'(d) - g(d)\}\}$ ;

[0078] 首先, 计算集合  $E_o^{LL}$  的一个子集:  $E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} \in D^h\}$ ,

[0079] 假设  $E_a = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_q^l, v_q^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_q^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_q^h \in V_2^{LL'}$ ; 计算  $\psi' = \arg \max_{v \in V} \{||N'(\psi(v))|| / L(\psi(v))\}$ , 其中  $V = \{v_1^l, v_2^l, \dots, v_q^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$

[0080] 然后, 计算集合  $E_a$  的一个子集,  $E_b = \{(v_i^l, v_i^h) \in E_a \mid \psi(v_i^l) = \psi'\}$ ,

[0081] 最后,从集合 $E_b$ 中随机选择一条边 $e$ ,并将该边从二分图 $G^{LL'}=(V'_1, V'_2, E^{LL'})$ 和 $G'=(V'_1, V'_2, E')$ 中删除,即更新 $E^{LL'} \leftarrow E^{LL'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,更新二分图 $G^{LL'}$ 中 $V'_2$ 的节点度分布 $\{g'(d) \mid d=1, 2, \dots, d_{\max}\}$ ,其中 $g'(d)$ 表示 $V'_2$ 中度为 $d$ 的节点比率,转到步骤S712;

[0082] S711,假设 $E_o^{LH} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ ,其中 $v_1^l, v_2^l, \dots, v_r^l \in V'_1$ 且 $v_1^h, v_2^h, \dots, v_r^h \in V'_2$ ;计算 $\psi' = \arg_{\psi(v)} \max_{v \in V} \{ ||N'(\psi(v))|| / L(\psi(v)) \}$ ,其中 $V = \{v_1^l, v_2^l, \dots, v_r^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V'_1 \wedge \psi(v_1) = \psi(v)\}$ ,

[0083] 计算集合 $E_o^{LH}$ 的一个子集: $E_b = \{(v_i^l, v_i^h) \in E_o^{LH} \mid \psi(v_i^l) = \psi'\}$ ,

[0084] 并从集合 $E_b$ 中随机选择一条边 $e$ ,并将该边从二分图 $G^{LH'}=(V'_1, V'_2, E^{LH'})$ 和 $G'=(V'_1, V'_2, E')$ 中删除,即更新 $E^{LH'} \leftarrow E^{LH'} / \{e\}$ 且 $E' \leftarrow E' / \{e\}$ ,转到步骤S712;

[0085] S712,更新 $t \leftarrow t+1$ ,转到步骤S705;

[0086] S713,输出:采样子图 $G'=(V'_1, V'_2, E')$ 。

[0087] 其中,对 $\forall w \in W$ ,集合 $N'(w) = \{v_1 \mid v_1 \in V'_1 \wedge \psi(v_1) = w\}$ ,其中 $W$ 为步骤S701的输入, $V'_1 \subseteq V_1$ ,S701至S713期望对 $\forall w \in W$ 满足 $||N'(w)||$ 接近于 $L(w)$ ;

[0088] 其中, $||S||$ 表示集合 $S$ 的势,即集合 $S$ 包含元素的总数; $x \leftarrow y$ 表示将右边的函数或数值 $y$ 赋值给左边的变量 $x$ ; $d_v$ 表示节点 $v$ 在图 $G'=(V'_1, V'_2, E')$ 中的度; $A/B$ :如果 $A$ 和 $B$ 为数值,则表示除法,如果 $A$ 和 $B$ 为集合,则表示从集合 $A$ 中删除集合 $B$ 的所有元素; $\arg_d \max_{d \in D} \{g'(d) - g(d)\}$ 表示 $g'(d) - g(d)$ 取最大值时参数 $d$ 的取值; $\psi' = \arg_{\psi(v)} \max_{v \in V} \{ ||N'(\psi(v))|| / L(\psi(v)) \}$ 表示将 $||N'(\psi(v))|| / L(\psi(v))$ 取最大值时参数 $\psi(v)$ 的取值赋值给 $\psi'$ ; $\in$ 表示“属于”; $\notin$ 表示“不属于”; $\wedge$ 表示“逻辑与”; $\cup$ 表示集合的并集; $\forall$ 表示“任意”;节点的度是指与该节点相邻边的总数;输入参数规模缩减率 $R_R$ 被定义为需要从输入的二分图中删除边的总数与输入的二分图包含边数的比率。

[0089] 本发明还提供了一种互联网测试床拓扑结构大比例规模缩减装置,所述装置包括:

[0090] 建模单元,用于获取真实世界互联网拓扑结构的探测数据,以简单无向图 $G=(V, E)$ 建模自治系统级互联网拓扑结构,其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域,集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径;

[0091] 分解拓扑单元,用于采用互联网拓扑结构模型分解互联网拓扑图 $G=(V, E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $QI_1-RI$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ ,一个匹配图 $BI-BI$ 和一个内核图 $Core$ ;

[0092] 二分图采样1单元,用于调用二分图采样方法1对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $QI_1-RI$ 进行采样,从而抽取部分节点和边构成采样子图;

[0093] 匹配图采样单元,用于调用匹配图采样方法对匹配图 $BI-BI$ 进行采样,抽取部分节点和边构成采样子图;

[0094] 二分图采样2单元,用于调用二分图采样方法2对二分图 $Q_4-RI_1$ 进行采样,抽取部分

节点和边构成采样子图；

[0095] 二分图采样3单元,用于调用二分图采样方法3对二分图QI-PI和二分图Q-P进行采样,抽取部分节点和边构成采样子图；

[0096] 二分图采样4单元,用于调用二分图采样方法4对二分图Q<sub>3</sub>-PI进行采样,抽取部分节点和边构成采样子图；

[0097] 子图合并单元,用于对得到的所有采样子图进行合并,得到互联网外围拓扑所有采样子图的合并图；

[0098] 内核图采样单元,用于调用内核图采样方法对内核图进行采样,从而抽取部分节点和边构成采样子图；

[0099] 采样子图获取单元,用于对得到的外围拓扑采样子图的合并图和内核拓扑的采样子图进行合并,从而得到真实世界互联网拓扑图的采样子图；

[0100] 仿真测试单元,依据得到的真实世界互联网拓扑图的规模缩减采样子图,配置互联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境。

[0101] 本公开的有益效果为:本发明公开了一种互联网测试床拓扑结构大比例规模缩减方法,采用面向真实世界互联网拓扑的结构模型,设计测试床的拓扑结构大比例规模缩减技术,在维持拓扑属性稳定不变的前提条件下大比例降低测试床构建所需的仿真资源节点数量,解决有限仿真资源条件下的测试床构建难点问题,从而达到大幅度降低测试床构建成本和测试任务运行时间复杂性等经济效益,本发明可以在保持节点度分布、聚类系数、RichClub连通性和路径长度等重要拓扑属性稳定,即近似保持不变的前提下,大比例地缩减互联网测试床的自治系统级拓扑规模,规模缩减率可以达到96%以上。

## 附图说明

[0102] 通过对结合附图所示出的实施方式进行详细说明,本公开的上述以及其他特征将更加明显,本公开附图中相同的参考标号表示相同或相似的元素,显而易见地,下面描述中的附图仅仅是本公开的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图,在附图中:

[0103] 图1所示为自治系统级互联网拓扑独有的二分图结构节点分类及它们之间的关系；

[0104] 图2所示为七个二分图、一个匹配图和一个内核图；

[0105] 图3所示为互联网测试床拓扑结构大比例规模缩减技术方案的流程图；

[0106] 图4所示为拓扑规模缩减测试床的组播路由协议测试需求。

## 具体实施方式

[0107] 以下将结合实施例和附图对本公开的构思、具体结构及产生的技术效果进行清楚、完整的描述,以充分地理解本公开的目的、方案和效果。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0108] 如图1所示为自治系统级互联网拓扑独有的二分图结构节点分类及它们之间的关系,图2所示为七个二分图、一个匹配图和一个内核图,图3所示为互联网测试床拓扑结构大

比例规模缩减技术方案的流程图,图4所示为拓扑规模缩减测试床的组播路由协议测试需求,下面结合图1、图2、图3与图4来阐述根据本公开的实施方式的一种互联网测试床拓扑结构大比例规模缩减方法。

[0109] 本公开提出一种互联网测试床拓扑结构大比例规模缩减方法,具体包括以下步骤:

[0110] S1. 获取真实世界互联网拓扑结构的探测数据,并以简单无向图 $G=(V,E)$ 建模自治系统级互联网拓扑结构,其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域,集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径。

[0111] S2. 采用互联网拓扑结构模型,分解互联网拓扑图 $G=(V,E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $Q_{I_1-RI}$ 、 $Q_4-RI_1$ 、 $Q_I-PI$ 、 $Q_3-PI$ 、 $Q-P$ ,一个匹配图 $BI-BI$ 和一个内核图 $Core$ 。分解方法的具体步骤详见第1.1节。

[0112] S3. 调用“二分图采样方法1”(具体步骤详见第1.2节),对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $Q_{I_1-RI}$ 进行采样,从而抽取部分节点和边构成采样子图,具体步骤分别见第1.3节、第1.4节和第1.6节。

[0113] S4. 调用“匹配图采样方法”,对匹配图 $BI-BI$ 进行采样,从而抽取部分节点和边构成采样子图,具体步骤详见第1.5节。

[0114] S5. 调用“二分图采样方法2”,对二分图 $Q_4-RI_1$ 进行采样,从而抽取部分节点和边构成采样子图,具体步骤详见第1.7节。

[0115] S6. 调用“二分图采样方法3”(具体步骤详见第1.8节),对二分图 $Q_I-PI$ 和二分图 $Q-P$ 进行采样,从而抽取部分节点和边构成采样子图,具体步骤分别见第1.9节和第1.12节。

[0116] S7. 调用“二分图采样方法4”(具体步骤详见第1.10节),对二分图 $Q_3-PI$ 进行采样,从而抽取部分节点和边构成采样子图,具体步骤详见第1.11节。

[0117] S8. 对S3至S7得到的所有采样子图进行合并,得到互联网外围拓扑所有采样子图的合并图,具体步骤详见第1.13节。

[0118] S9. 调用“内核图采样方法”,对内核图进行采样,从而抽取部分节点和边构成采样子图,具体步骤详见第1.14节。

[0119] S10. 对S8得到的外围拓扑采样子图的合并图和S9得到的内核拓扑的采样子图进行合并,从而得到真实世界互联网拓扑图的采样子图,具体步骤详见第1.15节。

[0120] S11. 依据S10得到的真实世界互联网拓扑图的规模缩减采样子图,配置互联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境。

[0121] 自治系统级互联网拓扑可以采用简单无向图 $G=(V,E)$ 进行建模,其中 $V$ 和 $E$ 分别为节点集和边集。本发明的目标是要得到一个规模大比例缩减(缩减率大于96%)的简单无向图 $G'$ ,要求 $G$ 和 $G'$ 在以下重要拓扑属性上保持相似;其中自治系统级互联网测试床是指该测试床的每个节点为真实互联网中一个自治系统(Autonomous System)的抽象。自治系统是指有权自主确定在本系统中应采用何种路由协议的小型单位。

[0122] 节点度分布二维属性 $\overline{F}(d)$  vs.  $d$ : 定义为 $\overline{F}(d) = \sum_{k>d} f(k)$ , 其中 $d$ 为节点度,  $\overline{F}(d)$ 为度大于 $d$ 的节点数与拓扑图节点总数的比率,  $f(k)$ 为拓扑图中度为 $k$ 节点的比率。二维属

性  $\overline{F}(d)$  vs.  $d$  在直角坐标系中, 是以  $d$  为变量且以  $\overline{F}(d)$  为函数值的二维函数曲线。

[0123] 聚类系数二维属性  $C(d)$  vs.  $d$ :

[0124] 定义为  $C(d) = 2T_d/d(d-1)$ , 其中  $T_d = \sum_{i=1,2,\dots,t} T(v_i)/t$ ,  $d$  为节点度,  $C(d)$  为节点度  $d$  的聚类系数,  $v_1, v_2, \dots, v_t$  为拓扑图中所有度为  $d$  的节点,  $t$  为拓扑图中度为  $d$  的节点总数,  $T(v_i)$  为节点  $v_i$  任意两个相邻节点之间边的总数。二维属性  $C(d)$  vs.  $d$  在直角坐标系中, 是以  $d$  为变量且以  $C(d)$  为函数值的二维函数曲线。

[0125] RichClub 连通性二维属性  $\rho(r)$  vs.  $r$ : 定义  $\rho(r)$  为拓扑图中前  $r$  个度最大的节点构成子图包含边的总数占据该子图中最大可能边数 (即,  $r(r-1)/2$ ) 的比率;  $r$  的取值范围是从 1 到  $n$  的全部自然数, 其中  $n$  为拓扑图包含的节点总数。二维属性  $\rho(r)$  vs.  $r$  在直角坐标系中, 是以  $r$  为变量且以  $\rho(r)$  为函数值的二维函数曲线。

[0126] 路径长度分布二维属性  $\mu(1)$  vs.  $1$ : 定义  $\mu(1)$  为拓扑图中相互间最短路径长度为 1 的节点对总数占据全部节点对总数的比率。二维属性  $\mu(1)$  vs.  $1$  在直角坐标系中, 是以 1 为变量且以  $\mu(1)$  为函数值的二维函数曲线。

[0127] 其中 vs. 表示“相对于”, 即表示 vs. 前的函数是相对于 vs. 后的变量的函数。例如,  $\overline{F}(d)$  vs.  $d$  表示  $\overline{F}(d)$  是相对于变量  $d$  的函数。

[0128] 为了大比例缩减自治系统级互联网拓扑的节点规模, 需要了解该拓扑独有的结构特征。本发明发现该拓扑模型可以被分解为八类节点集:

$$[0129] \left\{ \begin{array}{l} P = \{v \in V \mid d_v = 1\} \\ Q = \{v \in V \mid \exists w, (v, w) \in E, w \in P\} \\ PI = \{v \in V \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) > 1\} \\ QI = \{v \in V \mid \exists w, (v, w) \in E_I, w \in PI\} \\ RI = \{v \in V \mid d_I(v) \geq 2 \wedge \forall (v, w) \in E_I, w \in QI\} \\ BI = \{v \in V \mid d_I(v) = 1 \wedge \forall (v, w) \in E_I, d_I(w) = 1\} \\ II = \{v \in V \mid d_I(v) = 0\} \\ OI = V_I / (PI \cup QI \cup RI \cup BI \cup II) \end{array} \right. , \quad (1)$$

[0130] 模型中参数说明如下:

[0131]  $V, E$ : 分别为拓扑图  $G = (V, E)$  的节点集和边集;

[0132]  $d_v$ : 为拓扑图  $G = (V, E)$  中节点  $v$  的度;

[0133]  $G_I = (V_I, E_I)$ : 为拓扑图  $G = (V, E)$  中, 由节点集  $R(G) = V / (P \cup Q)$  和两个端点均属于  $R(G)$  的所有边构成边集, 生成的子图;

[0134]  $V_I, E_I$ : 分别为子图  $G_I = (V_I, E_I)$  的节点集和边集;

[0135]  $d_I(v)$ : 为子图  $G_I = (V_I, E_I)$  中节点  $v$  的度;

[0136]  $\in, \exists, \forall, \wedge$ : 分别表示“属于”、“存在”、“任意”和“逻辑与”;

[0137]  $\cup$ : 表示集合的并集;

[0138]  $A/B$ : 表示从集合  $A$  中删除集合  $B$  的所有元素。

[0139] 上述拓扑模型中  $OI$  表示系统噪声节点, 其在真实系统占据的比率通常不超过 1.9%。为了简化问题, 本发明的互联网拓扑规模缩减方法将忽略  $OI$  节点。

[0140] 进一步地,式(1)中七类节点之间的连接关系。如图1所示,每个P节点仅且必被连接至唯一的Q节点,每个II节点仅且必被连接至至少两个Q节点,每个PI节点仅且必被连接至唯一的QI节点和至少一个Q节点,每个BI节点仅且必被连接至另外一个BI节点和至少一个Q节点,每个RI节点仅且必被连接至至少两个QI节点,此外QI和RI节点可能被连接至Q节点。由上述连接关系可以将自治系统级互联网拓扑分解为七个二分图 $Q_1-II$ 、 $Q-P$ 、 $Q_2-BI$ 、 $Q_3-PI$ 、 $QI-PI$ 、 $QI_1-RI$ 和 $Q_4-RI_1$ ,其中节点集 $Q_1 = \{v \in Q | \exists w \in II, (v, w) \in E\}$ 、

$Q_2 = \{v \in Q | \exists w \in BI, (v, w) \in E\}$ 、 $Q_3 = \{v \in Q | \exists w \in PI, (v, w) \in E\}$ 、 $Q_4 = \{v \in Q | \exists w \in RI, (v, w) \in E\}$ 、

$QI_1 = \{v \in QI | \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI | \exists v \in Q, (v, w) \in E\}$ ,一个匹配图BI-BI和一个内核图Core,如图2所示。在图2中的箭头表示节点集之间的包含关系,其箭头起点的集合被包含于箭头终点的集合。其中,P节点是指属于集合P的节点,Q节点是指属于集合Q的节点,II节点是指属于集合II的节点,PI节点是指属于集合PI的节点,QI节点是指属于集合QI的节点,BI节点是指属于集合BI的节点,RI节点是指属于集合RI的节点。

[0141] 其中,二分图是指图中全部节点可以被划分至两个不同的节点集,且图中每条边的两个端点必定分别属于这两个不同的节点集。二分图 $Q_1-II$ 是指由节点集 $Q_1$ 、节点集II和两个端点分别属于 $Q_1$ 和II的所有边构成边集生成的二分图结构。二分图 $Q-P$ 、 $Q_2-BI$ 、 $Q_3-PI$ 、 $QI-PI$ 、 $QI_1-RI$ 和 $Q_4-RI_1$ 的物理意义解释方法与二分图 $Q_1-II$ 相同。匹配图BI-BI是由节点集BI和两个端点都属于BI的所有边构成边集生成的图结构。内核图Core是由节点集 $Q \cup QI$ 和连接该节点集内部节点的所有边构成边集生成的图结构。

[0142] 本发明的规模缩减技术,将依据上述自治系统级互联网拓扑独有的二分图结构展开设计。

[0143] 二分图定义:如果简单无向图 $G = (V, E)$ 的节点集 $V = V_1 \cup V_2$ ,其中 $V_1 \cap V_2 = \Phi$ 且对任意一条边 $e = (u, w) \in E$ 的两个端点u和w不可能同时属于 $V_1$ 或同时属于 $V_2$ ,则该图可以表示为 $G = (V_1, V_2, E)$ ,且该图被称为二分图。

[0144] 其中, $\cup$ 表示集合的并集; $\cap$ 表示集合的交集; $\Phi$ 表示不包含任何元素的空集; $V$ 和 $E$ 分别为图G的节点集和边集。

[0145] 其中,互联网外围拓扑是由互联网外围节点组成的拓扑,互联网内核拓扑是由互联网内核节点组成的拓扑;其中外围节点是指互联网中只进行网络数据的发送与接收,而不进行网络数据转发的节点,内核节点是指互联网中可进行网络数据转发的节点。互联网外围拓扑由其拓扑模型中七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $QI_1-RI$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ 和一个匹配图BI-BI组成,互联网内核图由其拓扑模型中内核图Core构成。

[0146] 1.1互联网拓扑图分解

[0147] 输入:原始互联网拓扑图 $G = (V, E)$ 。

[0148] 输出:七个二分图 $G_{Q_1-II}$ 、 $G_{Q_2-BI}$ 、 $G_{QI_1-RI}$ 、 $G_{Q_4-RI_1}$ 、 $G_{QI-PI}$ 、 $G_{Q_3-PI}$ 、 $G_{Q-P}$ ,一个匹配图 $G_{BI-BI}$ 和一个内核图 $G_{Core}$ 。

[0149] 方法的流程与步骤:

[0150] S1.依据式(1)计算得到图G的七个节点集P、Q、II、BI、QI、PI和RI。

[0151] S2.计算二分图 $Q_1-II$ : $G_{Q_1-II} = (II, Q_1, E_{Q_1-II})$ ,其中 $Q_1 = \{v \in Q | \exists w \in II, (v, w) \in E\}$ 且 $E_{Q_1-II} = \{e = (v, w) | e \in E \wedge v \in II \wedge w \in Q_1\}$ 。

[0152] S3. 计算二分图 $Q_2$ -BI:  $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$ , 其中  $Q_2 = \{v \in Q \mid \exists w \in BI, (v, w) \in E\}$  且  $E_{Q_2-BI} = \{e = (v, w) \mid e \in E \wedge v \in BI \wedge w \in Q_2\}$ 。

[0153] S4. 计算二分图 $QI_1$ -RI:  $G_{QI_1-RI} = (RI, QI_1, E_{QI_1-RI})$ , 其中  $QI_1 = \{v \in QI \mid \exists w \in RI, (v, w) \in E\}$  且  $E_{QI_1-RI} = \{e = (v, w) \mid e \in E \wedge v \in RI \wedge w \in QI_1\}$ 。

[0154] S5. 计算二分图 $Q_4$ -RI $_1$ :  $G_{Q_4-RI_1} = (RI_1, Q_4, E_{Q_4-RI_1})$ , 其中  $Q_4 = \{v \in Q \mid \exists w \in RI, (v, w) \in E\}$ 、 $RI_1 = \{w \in RI \mid \exists v \in Q, (v, w) \in E\}$  且  $E_{Q_4-RI_1} = \{e = (v, w) \mid e \in E \wedge v \in RI_1 \wedge w \in Q_4\}$ 。

[0155] S6. 计算二分图 $QI$ -PI:  $G_{QI-PI} = (QI, PI, E_{QI-PI})$ , 其中  $E_{QI-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in QI\}$ 。

[0156] S7. 计算二分图 $Q_3$ -PI:  $G_{Q_3-PI} = (PI, Q_3, E_{Q_3-PI})$ , 其中  $Q_3 = \{v \in Q \mid \exists w \in PI, (v, w) \in E\}$  且  $E_{Q_3-PI} = \{e = (v, w) \mid e \in E \wedge v \in PI \wedge w \in Q_3\}$ 。

[0157] S8. 计算二分图 $Q$ -P:  $G_{Q-P} = (Q, P, E_{Q-P})$ , 其中  $E_{Q-P} = \{e = (v, w) \mid e \in E \wedge v \in P \wedge w \in Q\}$ 。

[0158] S9. 计算匹配图 $BI$ -BI:  $G_{BI-BI} = (BI, E_{BI-BI})$ , 其中  $E_{BI-BI} = \{e = (v, w) \mid e \in E \wedge v, w \in BI\}$ 。

[0159] S10. 计算内核图Core:  $G_{Core} = (V_{Core}, E_{Core})$ , 其中  $V_{Core} = Q \cup QI$  且  $E_{Core} = \{e = (v, w) \mid e \in E \wedge v, w \in V_{Core}\}$ 。

[0160] 其中,  $\in$  表示“属于”;  $\exists$  表示“存在”;  $\wedge$  表示“逻辑与”;  $\cup$  表示集合的并集;  $V$  和  $E$  分别为互联网拓扑图 $G$ 的节点集和边集。

[0161] 1.2二分图采样方法1

[0162] 输入: 二分图 $G = (V_1, V_2, E)$ , 其中  $V_1, V_2$  为节点集、 $E$  为边集, 参数 $T, R$ , 规模缩减率 $R_R$ 。其中, 规模缩减率 $R_R$ 被定义为需要从输入(待采样)的二分图 $G = (V_1, V_2, E)$ 中删除边的总数与该二分图包含边数的比率。

[0163] 输出: 采样子图 $G' = (V'_1, V'_2, E')$ 。

[0164] 符号定义:  $\|S\|$  表示集合 $S$ 的势, 即集合 $S$ 包含元素的总数;

[0165]  $x \leftarrow y$  表示将右边的函数或数值 $y$ 赋值给左边的变量 $x$ ;

[0166]  $d_v$  表示节点 $v$ 在图 $G' = (V'_1, V'_2, E')$ 中的度;

[0167]  $\Phi$  表示不包含任何元素的空集;

[0168]  $A/B$ : 如果 $A$ 和 $B$ 为数值, 则表示除法, 如果 $A$ 和 $B$ 为集合, 则表示从集合 $A$ 中删除集合 $B$ 的所有元素;

[0169]  $d^h = \arg \max_{d \in D} \{g'(d) - g(d)\}$  表示将 $g'(d) - g(d)$ 取最大值时参数 $d$ 的取值赋值给 $d^h$ ;

[0170]  $\in$  表示“属于”;  $\notin$  表示“不属于”;  $\wedge$  表示“逻辑与”;  $\cup$  表示集合的并集;

[0171] 节点的度是指与该节点相邻边的总数;

[0172]  $\|S = \{v \mid v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}$  表示集合  $S = \{v \mid v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}$  的势, 即集合 $S$ 包含元素的总数。

[0173] 方法的流程与步骤:

[0174] S1. 将节点集 $V_2$ 中全部节点按度从大到小顺利排列,并抽取前 $n_h$ 个最大度节点构成节点子集 $V_2^{HH}$ ,其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ;也就是说, $n_h$ 可以由 $V_2$ 包含节点的度和输入参数 $T$ 唯一确定。转S2。

[0175] S2. 计算节点集 $V_2^{LL} = V_2/V_2^{HH}$ ,并将二分图 $G$ 分解为 $G^{HH} = (V_1, V_2^{HH}, E^{HH})$ 和 $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ ,其中边集 $E^{HH} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{HH}\}$ 且边集 $E^{LL} = \{(v_1, v_2) \in E \mid v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ 。转S3。

[0176] S3. 设 $d_1, d_2, \dots, d_s$ 为二分图 $G$ 中节点集 $V_1$ 的全部节点度,并设 $d_{\max}$ 为节点集 $V_2^{LL}$ 中的最大节点度。计算节点集 $V_2^{LL}$ 的度分布 $\{g(d) \mid d=1, 2, \dots, d_{\max}\}$ ,其中 $g(d)$ 表示 $V_2^{LL}$ 中度为 $d$ 的节点比率。初始化二分图 $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ ,初始化二分图 $G^{HH'} = (V'_1, V_2^{HH'}, E^{HH'}) \leftarrow G^{HH} = (V_1, V_2^{HH}, E^{HH})$ ,初始化二分图 $G^{LL'} = (V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL} = (V_1, V_2^{LL}, E^{LL})$ ,初始化二分图 $G^{LL'}$ 中节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) \mid d=1, 2, \dots, d_{\max}\} \leftarrow \{g(d) \mid d=1, 2, \dots, d_{\max}\}$ 。计算需删除边的总数 $E_d = ||E||_{R_R}$ 。初始化变量 $t=0$ 。转S4。

[0177] S4. 如果 $t < E_d$ ,转S5;否则 $G' = (V'_1, V'_2, E')$  (其中 $V'_2 = V_2^{HH'} \cup V_2^{LL'}$ 且 $E' = E^{HH'} \cup E^{LL'}$ )为最终输出结果。

[0178] S5. 如果 $\|S = \{v \mid v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} \| > 0$ ,任意选择集合 $S$ 中一个节点 $v$ ,并计算两个可选边集 $E_o^{HH} = \{(v, v_2) \in E^{HH'} \mid v_2 \in V_2^{HH'}\}$ 和 $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v_2 \in V_2^{LL'}\}$ 。如果 $\|E_o^{HH}\| = 0$ ,转S7;如果 $\|E_o^{LL}\| = 0$ ,转S8;否则(即 $\|E_o^{HH}\| > 0$ 且 $\|E_o^{LL}\| > 0$ ),以概率 $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$ 转S7,并以概率 $1 - P^{LL'}$ 转S8。

[0179] 其中, $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ , $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ , (2)

$$[0180] \quad m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R), \quad m^{LH} = E_d - m^{LL}。 (3)$$

[0181] S6. 如果 $S = \Phi$ ,以式(4)定义的离散概率分布 $p_k / \sum_{i=1, 2, \dots, s} p_i$  ( $k=1, 2, \dots, s$ )随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ ,其中 $d_1 > d_2 > \dots > d_s > 0$ , $f(d_1), f(d_2), \dots, f(d_s)$ 表示节点集 $V_1$ 中度为 $d_k$  ( $1 \leq k \leq s$ )的节点比率。

$$[0182] \quad p_k = \sum_{i=1}^k f(d_i), (4)$$

[0183] 计算两个可选边集 $E_o^{HH} = \{(v, v_2) \in E^{HH'} \mid v \in V'_1 \wedge v_2 \in V_2^{HH'} \wedge d_v = d\}$ ,

$E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V'_1 \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}$ 。如果 $\|E_o^{HH}\| = 0$ ,转S7;如果 $\|E_o^{LL}\| = 0$ ,转S8;否则,以式(2)和(3)定义的概率 $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$ 转S7,并以概率 $1 - P^{LL'}$ 转S8。

[0184] S7. 假设边集 $E_o^{LL} = \{(v'_1, v_1^h), (v'_2, v_2^h), \dots, (v'_r, v_r^h)\}$ ,其中 $v'_1, v'_2, \dots, v'_r \in V'_1$ 且 $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ ,计算 $d^h = \arg \max_{d \in D} \{g'(d) - g(d)\}$ ,其中 $D$ 为集合 $V_2^{LL'}$ 中节点 $v_1^h, v_2^h, \dots, v_r^h$ 的度组成的集合,并

计算边集 $E_o^{LL}$ 的一个子集

$$[0185] \quad E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} = d^h\}, \quad (5)$$

[0186] 从 $E_a$ 中均等概率地采样一条边 $e$ ,并将该边从二分图 $G^{LL'} = (V_1', V_2', E^{LL'})$ 和 $G' = (V_1', V_2', E')$ 中删除,即更新边集 $E^{LL'} \leftarrow E^{LL'} / \{e\}$ 和边集 $E' \leftarrow E' / \{e\}$ 。更新二分图 $G^{LL'}$ 中节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) \mid d=1, 2, \dots, d_{\max}\}$ ,其中 $g'(d)$ 表示 $V_2^{LL'}$ 中度为 $d$ 的节点比率。转S9。

[0187] S8.从边集 $E_o^{LH}$ 中均等概率分布地采样一条边 $e$ ,并将该边从二分图 $G^{LH'} = (V_1', V_2', E^{LH'})$ 和 $G' = (V_1', V_2', E')$ 中删除,即更新边集 $E^{LH'} \leftarrow E^{LH'} / \{e\}$ 和边集 $E' \leftarrow E' / \{e\}$ 。转S9。

[0188] S9.更新 $t \leftarrow t+1$ ,转S4。

[0189] 方法的应用效果:

[0190] 本节的二分图采样方法1可以在保证删除 $E_d = ||E|| \cdot R_R$ 条边的前提下,使得原始图 $G = (V_1, V_2, E)$ 和采样子图 $G' = (V_1', V_2', E')$ 趋向于拥有相同的节点度分布。其中节点度分布是指图中全部节点的度的概率分布。

[0191] 1.3二分图 $G_{Q-II} = (II, Q_1, E_{Q-II})$ 采样

[0192] 输入:第1.1节输出的二分图 $G_{Q-II} = (II, Q_1, E_{Q-II})$ ,预先设定参数 $T$ (缺省情况下赋值为1)和 $R$ (缺省值为0.096),规模缩减率 $R_R$ (缺省值为0.96)。

[0193] 其中缺省情况下参数的取值被称为缺省值,其是系统参数的默认值,即表示一个参数在被修改前的初始值。

[0194] 输出:采样二分图 $G'_{Q-II} = (II', Q_1', E'_{Q-II})$ 。

[0195] 方法的流程与步骤:

[0196] S1.建立与第1.2节“二分图采样方法1”输入的映射关系:

[0197] 将 $G_{Q-II} = (II, Q_1, E_{Q-II})$ 赋值于 $G = (V_1, V_2, E)$ ,其中 $V_1 = II, V_2 = Q_1, E = E_{Q-II}$ 。

[0198] S2.采用“二分图采样方法1”得到 $G_{Q-II}$ 的采样子图 $G'_{Q-II} = (II', Q_1', E'_{Q-II})$ 。

[0199] S3.如果采样子图 $G'_{Q-II} = (II', Q_1', E'_{Q-II})$ 的节点集 $II'$ 中存在度为1的节点,则将这些度为1的节点全部从采样子图 $G'_{Q-II}$ 中删除。

[0200] S4.  $G'_{Q-II}$ 为最终输出。

[0201] 1.4二分图 $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$ 采样

[0202] 输入:第1.1节输出的二分图 $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$ ,预先设定参数 $T$ (缺省情况下赋值为1)和 $R$ (缺省值为0.026),规模缩减率 $R_R$ (缺省值为0.96)。

[0203] 输出:采样二分图 $G'_{Q_2-BI} = (BI', Q_2', E'_{Q_2-BI})$ 。

[0204] 方法的流程与步骤:

[0205] S1.建立与第1.2节“二分图采样方法1”输入的映射关系:

[0206] 将  $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$  赋值于  $G = (V_1, V_2, E)$ , 其中  $V_1 = BI, V_2 = Q_2, E = E_{Q_2-BI}$ 。

[0207] S2. 调用“二分图采样方法1”得到  $G_{Q_2-BI}$  的采样子图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$ 。

[0208] S3. 如果采样子图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$  的节点集  $BI'$  中节点总数为奇数, 则在  $BI'$  中选择一个度最小的节点, 并将该节点从采样子图  $G'_{Q_2-BI}$  中删除。

[0209] S4.  $G'_{Q_2-BI}$  为最终输出。

[0210] 1.5 匹配图采样方法 (匹配图  $G_{BI-BI} = (BI, E_{BI-BI})$  采样)

[0211] 输入: 第1.1节输出的二分图  $G_{Q_2-BI} = (BI, Q_2, E_{Q_2-BI})$  和匹配图  $G_{BI-BI} = (BI, E_{BI-BI})$ , 第1.4节输出的采样二分图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$ 。

[0212] 输出: 采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$ 。

[0213] 符号定义:  $d_v^{BI}$  表示  $BI$  节点  $v$  在二分图  $G_{Q_2-BI}$  中的度,  $BI$  节点  $v$  是指属于集合  $BI$  的节点  $v$ 。

[0214]  $d_v^{BI'}$  表示  $BI'$  节点  $v$  在采样图  $G'_{Q_2-BI}$  中的度,  $BI'$  节点  $v$  是指属于集合  $BI'$  的节点  $v$ 。

[0215]  $||S||$  表示集合  $S$  的势, 即集合  $S$  包含元素的总数。

[0216]  $\Phi$  表示空集;  $\lfloor x \rfloor$  表示元素  $0$  到  $x$  之间最接近于  $x$  的整数。

[0217]  $x \leftarrow y$  表示将右边的函数或数值  $y$  赋值给左边的变量  $x$ 。

[0218]  $\max(x, y)$  表示取  $x$  和  $y$  中最大值;  $\min(x, y)$  表示取  $x$  和  $y$  中最小值。

[0219]  $\in$  表示“属于”;  $\wedge$  表示“逻辑与”;  $\forall$  表示“任意”;  $\cup$  表示集合的并集。

[0220] 方法的流程与步骤:

[0221] S1. 对每条边  $e = (u, v) \in E_{BI-BI}$  (其中  $u, v \in BI$ ), 生成一个节点度对  $(d_u^{BI}, d_v^{BI})$  (不失一般性, 假设  $d_u^{BI} \leq d_v^{BI}$ ), 并将节点度对  $(d_u^{BI}, d_v^{BI})$  映射为  $x = d_v^{BI}(d_v^{BI} - 1)/2 + d_u^{BI}$ 。计算集合  $X = \{x = d_v^{BI}(d_v^{BI} - 1)/2 + d_u^{BI} \mid (u, v) \in E_{BI-BI}\}$ , 并计算节点度对分布函数  $\rho(x) = ||\{y \mid y \in X \wedge y = x\}|| / ||X||$ , 其表示集合  $X$  中出现元素  $x$  的概率值。

[0222] S2. 设采样二分图  $G'_{Q_2-BI}$  中  $BI'$  节点的最大度为  $d_m$ 。初始化边集  $E'_{BI-BI} = \Phi$ 。将节点度  $d_1$  以步长  $1$  从  $1$  增加至  $d_m - 1$ , 并将节点度  $d_2$  以步长  $1$  从  $d_1$  增加至  $d_m$ : 对每个节点度对  $(d_1, d_2)$ , 首先计算其期望数量  $m_{d_1, d_2}$  和最大可能数量  $m'_{d_1, d_2}$ , 其中

$$[0223] \quad m_{d_1, d_2} = \frac{||BI'||}{2} \cdot \rho\left(\frac{d_2(d_2 - 1)}{2} + d_1\right), \quad (6)$$

$$[0224] \quad m'_{d_1, d_2} = \begin{cases} \max\left(\left\|\{v \in BI' \mid d_v^{BI'} = d_1\}\right\|, \left\|\{v \in BI' \mid d_v^{BI'} = d_2\}\right\|\right) & \text{如果 } d_1 < d_2 \\ \left\|\{v \in BI' \mid d_v^{BI'} = d_1\}\right\| / 2 & \text{如果 } d_1 = d_2 \end{cases}; \quad (7)$$

[0225] 然后,标准分布(概率均等)地随机生成  $s = \min(m_{d_1, d_2}, m'_{d_1, d_2})$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, s$ ), 要求对  $\forall 1 \leq i, j \leq s$  满足  $u_i \in BI' \wedge v_i \in BI' \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j \wedge d_{u_i}^{BI'} = d_1 \wedge d_{v_i}^{BI'} = d_2$ ; 最后,更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1,2,\dots,s}$ 。

[0226] S3. 对于节点集  $Y = \{v \in BI' \mid \forall (u, w) \in E'_{BI-BI}, u \neq v \wedge w \neq v\}$ , 标准分布地随机生成  $t = |BI'| \mid / 2 - |E'_{BI-BI}|$  条边  $e_i = (u_i, v_i)$  ( $i = 1, 2, \dots, t$ ), 要求对  $\forall 1 \leq i, j \leq t$  满足  $u_i \in Y \wedge v_i \in Y \wedge u_i \neq u_j \wedge u_i \neq v_j \wedge v_i \neq v_j$ , 并更新边集  $E'_{BI-BI} \leftarrow E'_{BI-BI} \cup \{e_i\}_{i=1,2,\dots,t}$ 。此时,生成的采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$  为本节的输出结果。

[0227] 1.6二分图  $G_{Q_1-R_1} = (RI, Q_1, E_{Q_1-R_1})$  采样

[0228] 输入:第1.1节输出的二分图  $G_{Q_1-R_1} = (RI, Q_1, E_{Q_1-R_1})$ , 预先设定参数T(缺省情况下赋值为1)和R(缺省值为0), 规模缩减率  $R_R$  (缺省值为0.96)。

[0229] 输出:采样二分图  $G'_{Q_1-R_1} = (RI', Q_1', E'_{Q_1-R_1})$ 。

[0230] 方法的流程与步骤:

[0231] S1. 建立与第1.2节“二分图采样方法1”输入的映射关系:

[0232] 将  $G_{Q_1-R_1} = (RI, Q_1, E_{Q_1-R_1})$  赋值于  $G = (V_1, V_2, E)$ , 其中  $V_1 = RI, V_2 = Q_1, E = E_{Q_1-R_1}$ 。

[0233] S2. 调用“二分图采样方法1”得到  $G_{Q_1-R_1}$  的采样子图  $G'_{Q_1-R_1} = (RI', Q_1', E'_{Q_1-R_1})$ 。

[0234] S3. 如果采样子图  $G'_{Q_1-R_1} = (RI', Q_1', E'_{Q_1-R_1})$  的节点集  $RI'$  中存在度为1的节点, 则将这些度为1的节点全部从采样子图  $G'_{Q_1-R_1}$  中删除。

[0235] S4.  $G'_{Q_1-R_1}$  为最终输出。

[0236] 1.7二分图采样方法2(二分图  $G_{Q_4-R_1} = (RI_1, Q_4, E_{Q_4-R_1})$  采样)

[0237] 输入:第1.1节输出的二分图  $G_{Q_4-R_1} = (RI_1, Q_4, E_{Q_4-R_1})$ , 第1.6节输出的采样二分图  $G'_{Q_1-R_1} = (RI', Q_1', E'_{Q_1-R_1})$ 。

[0238] 输出:采样二分图  $G'_{Q_4-R_1} = (RI'_1, Q'_4, E'_{Q_4-R_1})$ 。

[0239] 符号定义:  $\in$  表示“属于”;  $\wedge$  表示“逻辑与”;  $\exists$  表示“存在”。

[0240] 方法的流程与步骤:

[0241] S1. 计算节点集  $RI'_1 = \{w \in RI' \mid \exists v \in Q_4, (w, v) \in E_{Q_4-R_1}\}$ 、节点集

$Q'_4 = \{v \in Q_4 \mid \exists w \in RI', (w, v) \in E_{Q_4-R_1}\}$  和边集  $E'_{Q_4-R_1} = \{(w, v) \in E_{Q_4-R_1} \mid w \in RI'_1 \wedge v \in Q'_4\}$ 。

[0242] S2. 则,由节点集  $RI'_1$ 、节点集  $Q'_4$  和边集  $E'_{Q_4-R_1}$  生成的二分图  $G'_{Q_4-R_1} = (RI'_1, Q'_4, E'_{Q_4-R_1})$ , 即为最终的输出结果。

[0243] 1.8二分图采样方法3

[0244] 输入:二分图 $G=(V_1, V_2, E)$ , 其中 $V_1, V_2$ 为节点集、 $E$ 为边集, 节点集 $\bar{V}_1 \subseteq V_1$ , 参数 $T$ , 规模缩减率 $R_R$ 。其中, 规模缩减率 $R_R$ 被定义为需要从输入(待采样)的二分图 $G=(V_1, V_2, E)$ 中删除边的总数与该二分图包含边数的比率。

[0245] 输出:采样子图 $G'=(V'_1, V'_2, E')$ 。

[0246] 约束条件:节点集 $\bar{V}_1$ 要求被包含于采样子图, 即 $\bar{V}_1 \subseteq V'_1$ 。

[0247] 二分图 $G$ 中节点集 $V_2$ 包含节点的度均为1。

[0248] 采样子图 $G'$ 中节点集 $V'_2$ 包含节点的度均为1。

[0249] 符号定义: $||S||$ 表示集合 $S$ 的势, 即集合 $S$ 包含元素的总数。

[0250]  $x \leftarrow y$ 表示将右边的函数或数值 $y$ 赋值给左边的变量 $x$ 。

[0251] 当节点 $v \in V_1$ 或 $v \in V_1^H$ 或 $v \in V_1^L$ (其中 $V_1^H$ 和 $V_1^L$ 将在 $S1, S2$ 中定义),  $d_v$ 表示 $v$ 在图 $G=(V_1, V_2, E)$ 的度。其他情况或没有特定约束下,  $d_v$ 表示 $v$ 在图 $G'=(V'_1, V'_2, E')$ 的度。 $\in$ 表示“属于”;  $\notin$ 表示“不属于”;  $\wedge$ 表示“逻辑与”;  $\vee$ 表示“逻辑或”。

[0252]  $A/B$ :如果 $A$ 和 $B$ 为数值, 则表示除法, 如果 $A$ 和 $B$ 为集合, 则表示从集合 $A$ 中删除集合 $B$ 的所有元素。

[0253]  $||S=\{v|v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} / \tilde{V}_1||$ 表示集合 $S=\{v|v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} / \tilde{V}_1$

的势, 即集合 $S$ 包含元素的总数。

[0254] 方法的流程与步骤:

[0255] S1. 将节点集 $V_1$ 中全部节点按度从大到小顺利排列, 并抽取前 $n_h$ 个最大度节点构成节点子集 $V_1^H$ , 其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ; 也就是说, 参数 $n_h$ 可以由节点集 $V_1$ 包含全部节点的度和输入参数 $T$ 唯一确定。转S2。

[0256] S2. 计算节点集 $V_1^L=V_1/V_1^H$ , 设 $d_1, d_2, \dots, d_s$ 为二分图 $G$ 中节点集 $V_1$ 的全部节点度。初始化二分图 $G'=(V'_1, V'_2, E') \leftarrow G=(V_1, V_2, E)$ , 初始化节点集 $V_1^{L'} \leftarrow V_1^L$ , 初始化节点集 $V_1^{H'} \leftarrow V_1^H$ 。计算最终需要删除边的总数 $E_d=||E|| \cdot R_R$ 。初始化变量 $t=0$ 。转S3。

[0257] S3. 如果 $t < E_d$ , 则以概率 $P=\gamma^{L'} / (\gamma^{H'} + \gamma^{L'})$ 转S4, 并以概率 $1-P$ 转S7; 否则,  $G'=(V'_1, V'_2, E')$ (其中 $V'_1=V_1^{H'} \cup V_1^{L'}$ )为最终输出结果。其中,

$$[0258] \quad \gamma^{L'} = m^L - \left( \sum_{v \in V_1^L} d_v - \sum_{v \in V_1^{L'}} d_v \right), \quad \gamma^{H'} = m^H - \left( \sum_{v \in V_1^H} d_v - \sum_{v \in V_1^{H'}} d_v \right), \quad (8)$$

$$[0259] \quad m^L = \frac{E_d \cdot \sum_{v \in V_1^L} d_v}{\sum_{v \in V_1^H \cup V_1^L} d_v} \text{ 且 } m^H = E_d - m^L. \quad (9)$$

[0260] S4. 如果 $||S=\{v|v \in V_1^{L'} \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\} / \tilde{V}_1|| > 0$ , 其中节点集 $\tilde{V}_1 = \{v|d_v = 1 \wedge v \in \bar{V}_1\}$ , 则随机选择集合 $S$ 中一个节点 $v$ , 转S6; 否则, 转S5。

[0261] S5. 以式(10)定义的离散概率分布 $p_k / \sum_{i=1,2,\dots,s} p_i$  ( $k=1, 2, \dots, s$ ) 随机采样一个节点度 $d \in \{d_1, d_2, \dots, d_s\}$ , 其中 $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$ 表示节点集 $V_1$ 在二分图 $G=(V_1, V_2, E)$ 中度为 $d_k$  ( $1 \leq k \leq s$ )的节点比率。

$$[0262] \quad p_k = \sum_{i=1}^k f(d_i), \quad (10)$$

[0263] 计算两个节点集  $V^1 = \{v | v \in V_1^{L'} \cap \bar{V}_1 \wedge d_v = d\}$  和  $V^2 = \{v | v \in V_1^{L'} / \bar{V}_1 \wedge d_v = d\}$ 。如果  $||V^2|| > 0$ , 则随机选择集合  $V^2$  中一个节点  $v$ , 否则随机选择集合  $V^1$  中一个节点  $v$ ; 转S6。

[0264] S6. 如果  $d_v \geq 2 \vee v \notin \bar{V}_1$ , 则随机选择一条与节点  $v$  相邻的边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E')$ :  $E' \leftarrow E' / \{e\}$ ; 如果  $d_v = 0$ , 则更新节点集  $V_1^{L'} \leftarrow V_1^{L'} / \{v\}$ 。如果  $d_v < 2 \wedge v \in \bar{V}_1$ , 则此步不作任何操作。转S8。

[0265] S7. 从集合  $\{(w, v) \in E' | w \in V_2' \wedge v \in V_1^{H'}\} / \bar{V}_1$  中随机选择一条边  $e$ , 并更新二分图  $G' = (V'_1, V'_2, E')$ :  $E' \leftarrow E' / \{e\}$ ; 如果  $d_v = 0$ , 则更新节点集  $V_1^{H'} \leftarrow V_1^{H'} / \{v\}$ 。转S8。

[0266] S8. 更新  $t \leftarrow t + 1$ , 转S3。

[0267] 方法的应用效果:

[0268] 本节的二分图采样方法3可以在保证删除  $E_d = ||E|| \cdot R_R$  条边的前提下, 使得原始图  $G = (V_1, V_2, E)$  和采样子图  $G' = (V'_1, V'_2, E')$  趋向于拥有相同的节点度分布, 并要求采样子图  $G' = (V'_1, V'_2, E')$  保证满足  $\bar{V}_1 \subseteq V_1'$ 。其中节点度分布是指图中全部节点的度的概率分布。

[0269] 1.9二分图  $G_{QI-PI} = (QI, PI, E_{QI-PI})$  采样

[0270] 输入: 第1.1节输出的二分图  $G_{QI-PI} = (QI, PI, E_{QI-PI})$ , 第1.6节输出的采样二分图  $G'_{QI-RI} = (RI', QI', E'_{QI-RI})$ , 预先设定参数  $T$  (缺省情况下赋值为0), 规模缩减率  $R_R$  (缺省情况下赋值为0.96)。

[0271] 输出: 采样二分图  $G'_{QI-PI} = (QI', PI'', E'_{QI-PI})$ 。

[0272] 方法的流程与步骤:

[0273] S1. 建立与第1.8节“二分图采样方法3”输入  $G = (V_1, V_2, E)$  和  $\bar{V}_1$  的映射关系:

[0274] 将  $G_{QI-PI} = (QI, PI, E_{QI-PI})$  赋值于  $G = (V_1, V_2, E)$ , 其中  $V_1 = QI, V_2 = PI, E = E_{QI-PI}$ ; 并设节点集  $\bar{V}_1 = QI'$ 。

[0275] S2. 调用“二分图采样方法3”得到  $G_{QI-PI}$  的采样子图  $G'_{QI-PI} = (QI', PI'', E'_{QI-PI})$ 。

[0276] S3. 则,  $G'_{QI-PI}$  为最终输出。

[0277] 1.10二分图采样方法4

[0278] 输入: 二分图  $G = (V_1, V_2, E)$ , 其中  $V_1, V_2$  为节点集、 $E$  为边集, 多对一映射函数  $\psi: V_1 \rightarrow W$  (设  $\psi(v_1)$  为节点集  $W$  中与节点  $v_1 \in V_1$  对应的唯一节点), 对任意节点  $w \in W$  预选设定参数  $L(w)$ , 预选给定节点度分布  $\{g(d) | d = 1, 2, \dots\}$ , 预先设定参数  $T, R$ , 规模缩减率  $R_R$ 。其中, 规模缩减率  $R_R$  被定义为需要从输入(待采样)的二分图  $G = (V_1, V_2, E)$  中删除边的总数与该二分图包含边数的比率。

[0279] 输出: 采样子图  $G' = (V'_1, V'_2, E')$ 。

[0280] 对  $\forall w \in W$ , 定义集合  $N'(w) = \{v_1 | v_1 \in V'_1 \wedge \psi(v_1) = w\}$ , 其中  $V'_1 \subseteq V_1$ 。

[0281] 本节的二分图采样方法4期望对  $\forall w \in W$  满足  $||N'(w)||$  接近于  $L(w)$ 。

[0282] 符号定义:  $||S||$  表示集合  $S$  的势, 即集合  $S$  包含元素的总数。

[0283]  $x \leftarrow y$ 表示将右边的函数或数值 $y$ 赋值给左边的变量 $x$ 。

[0284]  $d_v$ 表示节点 $v$ 在图 $G' = (V'_1, V'_2, E')$ 中的度。

[0285]  $\in$ 表示“属于”； $\notin$ 表示“不属于”； $\wedge$ 表示“逻辑与”； $\forall$ 表示“任意”。

[0286]  $\cup$ 表示集合的并集。

[0287]  $A/B$ :如果 $A$ 和 $B$ 为数值,则表示除法,如果 $A$ 和 $B$ 为集合,则表示从集合 $A$ 中删除集合 $B$ 的所有元素。

[0288]  $\arg \max_{d \in \mathbb{D}} \{g'(d) - g(d)\}$ 表示 $g'(d) - g(d)$ 取最大值时参数 $d$ 的取值。

[0289]  $\psi' = \arg \psi(v) \max_{v \in V} \{ \|N'(\psi(v))\| / L(\psi(v)) \}$ 表示将 $\|N'(\psi(v))\| / L(\psi(v))$ 取最大值时参数 $\psi(v)$ 的取值赋值给 $\psi'$ 。

[0290] 节点的度是指与该节点相邻边的总数。

[0291]  $\|S = \{v | v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}$ 表示集合 $S = \{v | v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}$ 的势,即集合 $S$ 包含元素的总数。

[0292] 方法的流程与步骤:

[0293] S1.将节点集 $V_2$ 中全部节点按度从大到小顺利排列,并抽取前 $n_h$ 个最大度节点构成节点子集 $V_2^{HH}$ ,其中这 $n_h$ 个节点中相同度重复数的最大值为 $T$ ;也就是说,参数 $n_h$ 可以由节点集 $V_2$ 包含全部节点的度和输入参数 $T$ 唯一确定。转S2。

[0294] S2.计算节点集 $V_2^{LL} = V_2 / V_2^{HH}$ ,并将二分图 $G$ 分解为 $G^{HH} = (V_1, V_2^{HH}, E^{HH})$ 和 $G^{LL} = (V_1, V_2^{LL}, E^{LL})$ ,其中边集 $E^{HH} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{HH}\}$ 且边集 $E^{LL} = \{(v_1, v_2) \in E | v_1 \in V_1 \wedge v_2 \in V_2^{LL}\}$ 。

[0295] S3.设 $d_1, d_2, \dots, d_s$ 为二分图 $G$ 中节点集 $V_1$ 的全部节点度,并设 $d_{\max}$ 为节点集 $V_2^{LL}$ 中的最大节点度。初始化二分图 $G' = (V'_1, V'_2, E') \leftarrow G = (V_1, V_2, E)$ ,初始化二分图 $G^{HH'} = (V'_1, V_2^{HH'}, E^{HH'}) \leftarrow G^{HH} = (V_1, V_2^{HH}, E^{HH})$ ,初始化二分图 $G^{LL'} = (V'_1, V_2^{LL'}, E^{LL'}) \leftarrow G^{LL} = (V_1, V_2^{LL}, E^{LL})$ 。

计算节点集 $V_2^{LL'}$ 的度分布 $\{g'(d) | d = 1, 2, \dots, d_{\max}\}$ ,其中 $g'(d)$ 表示节点集 $V_2^{LL'}$ 中度为 $d$ 的节点比率。计算最终需要删除边的总数 $E_d = \|E\|_{R_R}$ 。初始化变量 $t = 0, \eta = 0$ 。转S4。

[0296] S4.如果 $t < E_d$ ,转S5;否则 $G' = (V'_1, V'_2, E')$ (其中 $V'_2 = V_2^{HH'} \cup V_2^{LL'}$ 且 $E' = E^{HH'} \cup E^{LL'}$ )为最终输出结果。

[0297] S5.如果 $\eta > 0 \wedge \|\tilde{V}_1\| > 0$ ,则更新 $\eta \leftarrow \eta - 1$ ,计算两个可选边集 $E_o^{HH'} = \{(v, v_2) \in E^{HH'} | v \in \tilde{V}_1 \wedge v_2 \in V_2^{HH'}\}$ 和 $E_o^{LL'} = \{(v, v_2) \in E^{LL'} | v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ ,其中

$$[0298] \quad \tilde{V}_1 = \left\{ v_1 \in V'_1 \mid \frac{\|N'(\psi(v_1))\|}{L(\psi(v_1))} > 1 \wedge d_v = 1 \right\}, \quad (11)$$

[0299] 且 $N'(\psi(v_1)) = \{v | v \in V'_1 \wedge \psi(v) = \psi(v_1)\}$ ,并转S8;否则,转S6。

[0300] S6.如果 $\|S = \{v | v \in V'_1 \wedge d_v \notin \{d_1, d_2, \dots, d_s\}\}\| > 0$ ,选择集合 $S$ 中一个节点 $v$ ,设该节点的度为 $d = d_v$ ;否则以式(12)定义的离散概率分布 $p_k / \sum_{i=1,2,\dots,s} p_i$  ( $k = 1, 2, \dots, s$ )随机采样一

个节点度  $d \in \{d_1, d_2, \dots, d_s\}$ , 其中  $d_1 > d_2 > \dots > d_s > 0$ ,  $f(d_1), f(d_2), \dots, f(d_s)$  表示节点集  $V_1$  中度为  $d_k (1 \leq k \leq s)$  的节点比率。转S7。

$$[0301] \quad p_k = \sum_{i=1}^k f(d_i), \quad (12)$$

[0302] S7. 如果  $d = 1 \wedge \|\tilde{V}_1\| = 0$ , 则更新  $\eta \leftarrow \eta + 1$ , 转S11; 如果  $d = 1 \wedge \|\tilde{V}_1\| > 0$ , 则计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LH'}\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in \tilde{V}_1 \wedge v_2 \in V_2^{LL'}\}$ , 转S8; 如果  $d > 1$ , 则计算两个可选边集  $E_o^{LH} = \{(v, v_2) \in E^{LH'} \mid v \in V_1' \wedge v_2 \in V_2^{LH'} \wedge d_v = d\}$  和  $E_o^{LL} = \{(v, v_2) \in E^{LL'} \mid v \in V_1' \wedge v_2 \in V_2^{LL'} \wedge d_v = d\}$ , 转S8。

[0303] S8. 如果  $\|E_o^{LH}\| = 0$ , 转S9; 如果  $\|E_o^{LL}\| = 0$ , 转S10; 如果  $\|E_o^{LH}\| > 0 \wedge \|E_o^{LL}\| > 0$ , 以概率  $P^{LL'} = \gamma^{LL'} / (\gamma^{LH'} + \gamma^{LL'})$  转S9, 以概率  $1 - P^{LL'}$  转S10。

[0304] 其中,  $\gamma^{LL'} = m^{LL} - (||E^{LL}|| - ||E^{LL'}||)$ ,  $\gamma^{LH'} = m^{LH} - (||E^{LH}|| - ||E^{LH'}||)$ , (13)

$$[0305] \quad m^{LL} = \frac{E_d \cdot \|E^{LL}\|}{\|E^{LH} \cup E^{LL}\|} \cdot (1 + R), \quad m^{LH} = E_d - m^{LL}. \quad (14)$$

[0306] S9. 假设  $E_o^{LL} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LL'}$ ; 计算  $D^h = \{d \in D \mid g'(d) - g(d) > 1/\|V_2^{LL'}\|\}$ , 其中  $D$  为集合  $V_2^{LL'}$  中节点  $v_1^h, v_2^h, \dots, v_r^h$  的度组成的集合; 如果  $||D^h|| = 0$ , 则更新  $D^h \leftarrow \{\arg \max_{d \in D} \{g'(d) - g(d)\}\}$ 。

[0307] 首先, 计算集合  $E_o^{LL}$  的一个子集:

$$[0308] \quad E_a = \{(v_i^l, v_i^h) \in E_o^{LL} \mid d_{v_i^h} \in D^h\}, \quad (15)$$

[0309] 假设  $E_a = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_q^l, v_q^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_q^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_q^h \in V_2^{LL'}$ ; 计算  $\psi' = \arg \max_{\psi(v)} \{||N'(\psi(v))|| / L(\psi(v))\}$ , 其中  $V = \{v_1^l, v_2^l, \dots, v_q^l\}$ ,  $N'(\psi(v)) = \{v_1 \mid v_1 \in V_1' \wedge \psi(v_1) = \psi(v)\}$ 。

[0310] 然后, 计算集合  $E_a$  的一个子集:

$$[0311] \quad E_b = \{(v_i^l, v_i^h) \in E_a \mid \psi(v_i^l) = \psi'\}, \quad (16)$$

[0312] 最后, 从集合  $E_b$  中随机选择一条边  $e$ , 并将该边从二分图  $G^{LL'} = (V_1', V_2^{LL'}, E^{LL'})$  和  $G' = (V_1', V_2', E')$  中删除, 即更新边集  $E^{LL'} \leftarrow E^{LL'} / \{e\}$  和边集  $E' \leftarrow E' / \{e\}$ 。更新二分图  $G^{LL'}$  中节点集  $V_2^{LL'}$  的节点度分布  $\{g'(d) \mid d = 1, 2, \dots, d_{\max}\}$ , 其中  $g'(d)$  表示  $V_2^{LL'}$  中度为  $d$  的节点比率。转S11。

[0313] S10. 假设  $E_o^{LH} = \{(v_1^l, v_1^h), (v_2^l, v_2^h), \dots, (v_r^l, v_r^h)\}$ , 其中  $v_1^l, v_2^l, \dots, v_r^l \in V_1'$  且  $v_1^h, v_2^h, \dots, v_r^h \in V_2^{LH'}$ ;

计算  $\psi' = \arg_{\psi(v)} \max_{v \in V} \{ ||N'(\psi(v))|| / L(\psi(v)) \}$ , 其中  $V = \{v'_1, v'_2, \dots, v'_r\}$ ,  $N'(\psi(v)) = \{v_1 | v_1 \in V'_1 \wedge \psi(v_1) = \psi(v)\}$ 。

[0314] 计算集合  $E_o^{LH}$  的一个子集:

$$[0315] \quad E_b = \{(v'_i, v'_i) \in E_o^{LH} | \psi(v'_i) = \psi'\}, \quad (17)$$

[0316] 并从集合  $E_b$  中随机选择一条边  $e$ , 并将该边从二分图  $G^{H'} = (V'_1, V'_2, E^{H'})$  和  $G' = (V'_1, V'_2, E')$  中删除, 即更新边集  $E^{LH'} \leftarrow E^{LH'} / \{e\}$  和边集  $E' \leftarrow E' / \{e\}$ 。转S11。

[0317] S11. 更新  $t \leftarrow t+1$ , 转S4。

[0318] 方法的应用效果:

[0319] 本节的二分图采样方法4可以在保证删除  $E_d = ||E|| \cdot R_R$  条边的前提下, 使得原始图  $G = (V_1, V_2, E)$  和采样子图  $G' = (V'_1, V'_2, E')$  趋向于拥有相同的节点度分布, 并要求采样子图  $G' = (V'_1, V'_2, E')$  保证对  $\forall w \in W$  满足  $||N'(w)||$  接近于  $L(w)$ 。其中节点度分布是指图中全部节点的度的概率分布。

[0320] 在第1.2节二分图采样方法1基础上设计本节二分图采样方法4的目标, 是为了保证图2中二分图QI-PI和二分图Q3-PI拥有相同的PI节点集。

[0321] 1.11二分图  $G_{Q_3-PI} = (PI, Q_3, E_{Q_3-PI})$  采样

[0322] 输入: 第1.1节输出的二分图  $G_{QI-PI} = (QI, PI, E_{QI-PI})$  和  $G_{Q_3-PI} = (PI, Q_3, E_{Q_3-PI})$ , 第1.9节输出的采样二分图  $G''_{QI-PI} = (QI', PI'', E''_{QI-PI})$ , 预先设定参数T (缺省情况下赋值为1)、R (缺省情况下赋值为1), 规模缩减率  $R_R$  (缺省情况下赋值为0.96)。

[0323] 输出: 采样二分图  $G'_{Q_3-PI} = (PI', Q'_3, E'_{Q_3-PI})$  和  $G'_{QI-PI} = (QI', PI', E'_{QI-PI})$ 。

[0324] 符号定义:  $N_v(G)$  表示节点  $v$  在图  $G$  的相邻节点集。

[0325]  $d_v(G)$  表示节点  $v$  在图  $G$  的度。

[0326]  $||S||$  表示集合  $S$  的势, 即集合  $S$  包含元素的总数。

[0327]  $\in$  表示“属于”;  $\exists$  表示“存在”;  $\wedge$  表示“逻辑与”;  $\cup$  表示集合的并集。

[0328] 方法的流程与步骤:

[0329] S1. 计算节点集  $\overline{PI} = \bigcup_{v \in QI'} N_v(G_{QI-PI}) = \{w \in PI | \exists v \in QI', (w, v) \in E_{QI-PI}\}$ 、节点集

$Q'_3 = \{v \in Q_3 | \exists w \in \overline{PI}, (w, v) \in E_{Q_3-PI}\}$  和边集  $E''_{Q_3-PI} = \{(w, v) \in E_{Q_3-PI} | w \in \overline{PI} \wedge v \in Q'_3\}$ , 定义由这两个

节点集和一个边集生成的二分图为  $G''_{Q_3-PI} = \{\overline{PI}, Q'_3, E''_{Q_3-PI}\}$ 。

[0330] S2. 建立与第1.10节“二分图采样方法4”输入的映射关系:

[0331] 将  $G''_{Q_3-PI} = \{\overline{PI}, Q'_3, E''_{Q_3-PI}\}$  赋值于  $G = (V_1, V_2, E)$ , 其中  $V_1 = \overline{PI}$ ,  $V_2 = Q'_3$ ,  $E = E''_{Q_3-PI}$ ; 设

定多对一映射函数  $\psi: \overline{PI} \rightarrow QI'$  为: 对  $v \in \overline{PI}$ ,  $\psi(v) \in QI' \wedge (v, \psi(v)) \in E_{QI-PI}$ ; 对  $\forall w \in QI'$  设定

预先参数  $L(w) = d_w(G''_{QI-PI})$ ; 设定  $\{g(d) | d = 1, 2, \dots\}$  为二分图  $G_{Q_3-PI} = \{PI, Q_3, E_{Q_3-PI}\}$  中节点

集  $Q_3$  包含节点的度分布, 其中  $g(d)$  表示度为  $d$  的节点比率。

[0332] S3. 调用二分图采样方法4对  $G'_{Q_3-PI}$  删除  $m_o \cdot R_{R-} (m_o - m''_o)$  条边, 其中  $m_o = \|E_{Q_3-PI}\|$  且  $m''_o = \|E'_{Q_3-PI}\|$ , 并生成采样结果图  $G'_{Q_3-PI} = (PI', Q'_3, E'_{Q_3-PI})$ 。

[0333] S4. 为了保证采样结果  $G'_{Q_3-PI} = (PI', Q'_3, E'_{Q_3-PI})$  和二分图  $G''_{QI-PI} = (QI', PI'', E''_{QI-PI})$  拥有相同的PI节点集, 将二分图  $G''_{QI-PI} = (QI', PI'', E''_{QI-PI})$  更新为二分图  $G'_{QI-PI} = (QI', PI', E'_{QI-PI})$ , 其中边集  $E'_{QI-PI} = \{(w, v) \in E_{QI-PI} | w \in QI' \wedge v \in PI'\}$ 。

[0334] S5. 则,  $G'_{Q_3-PI}$  和  $G'_{QI-PI}$  为最终的输出结果。

[0335] 1.12 二分图  $G_{Q-P} = (Q, P, E_{Q-P})$  采样

[0336] 输入: 第1.1节输出的二分图  $G_{Q-P} = (Q, P, E_{Q-P})$ , 第1.3节输出的采样二分图  $G'_{Q-I} = (I', Q'_1, E'_{Q-I})$ , 第1.4节输出的采样二分图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$ , 第1.7节输出的采样二分图  $G'_{Q_4-R_1} = (R_1', Q'_4, E'_{Q_4-R_1})$ , 第1.11节输出的采样二分图  $G'_{Q_3-PI} = (PI', Q'_3, E'_{Q_3-PI})$ 。参数 T (缺省值为3), 规模缩减率  $R_R$  (缺省值为0.96)。

[0337] 输出: 采样二分图  $G'_{Q-P} = (Q', P', E'_{Q-P})$  和  $Q'_1, Q'_2, Q'_3, Q'_4$  的分解节点集  $Q'_{11}, Q'_{21}, Q'_{31}, Q'_{41}, Q'_{12}, Q'_{22}, Q'_{32}, Q'_{42}$ 。

[0338] 符号定义:  $d_v(G)$  表示节点  $v$  在图  $G$  的度。

[0339]  $\cup$  表示集合的并集。

[0340]  $A/B$ : 如果  $A$  和  $B$  为集合, 则表示从集合  $A$  中删除集合  $B$  的所有元素。

[0341] 方法的流程与步骤:

[0342] S1. 分解节点集  $Q'_1, Q'_2, Q'_3$  和  $Q'_4$  如下:

$$[0343] \begin{cases} Q'_{11} = \{w \in Q'_1 | d_w(G'_{Q-I}) \geq 2\}, & Q'_{12} = Q'_1 / Q'_{11} \\ Q'_{21} = \{w \in Q'_2 | d_w(G'_{Q_2-BI}) \geq 2\}, & Q'_{22} = Q'_2 / Q'_{21} \\ Q'_{31} = \{w \in Q'_3 | d_w(G'_{Q_3-PI}) \geq 2\}, & Q'_{32} = Q'_3 / Q'_{31} \\ Q'_{41} = \{w \in Q'_4 | d_w(G'_{Q_4-R_1}) \geq 2\}, & Q'_{42} = Q'_4 / Q'_{41} \end{cases}, \quad (18)$$

[0344] S2. 建立与第1.8节“二分图采样方法3”输入  $G = (V_1, V_2, E)$  和  $\bar{V}_1$  的映射关系:

[0345] 将  $G_{Q-P} = (Q, P, E_{Q-P})$  赋值于  $G = (V_1, V_2, E)$ , 其中  $V_1 = Q, V_2 = P, E = E_{Q-P}$ 。并设节点集  $\bar{V}_1 = Q'_{11} \cup Q'_{21} \cup Q'_{31} \cup Q'_{41}$ 。

[0346] S3. 调用二分图采样方法3得到  $G_{Q-P}$  的采样子图  $G'_{Q-P} = (Q', P', E'_{Q-P})$ 。

[0347] S4. 则,  $G'_{Q-P}$  和式 (18) 的八个分解节点集  $Q'_{11}, Q'_{21}, Q'_{31}, Q'_{41}, Q'_{12}, Q'_{22}, Q'_{32}, Q'_{42}$  为最终输出。

[0348] 1.13 互联网外围拓扑所有采样子图的合并

[0349] 输入: 第1.3节输出的采样二分图  $G'_{Q-I} = (I', Q'_1, E'_{Q-I})$ , 第1.4节输出的采样二分图  $G'_{Q_2-BI} = (BI', Q'_2, E'_{Q_2-BI})$ , 第1.5节输出的采样匹配图  $G'_{BI-BI} = (BI', E'_{BI-BI})$ , 第1.6节输出的采样二分图  $G'_{Q_4-R_1} = (R_1', Q'_4, E'_{Q_4-R_1})$ , 第1.7节输出的采样二分图  $G'_{Q_3-PI} = (PI', Q'_3, E'_{Q_3-PI})$ ,

第1.11节输出的采样二分图  $G'_{Q_3-PI} = (PI', Q_3', E'_{Q_3-PI})$  和  $G'_{QI-PI} = (QI', PI', E'_{QI-PI})$ , 第1.12节输出的采样二分图  $G'_{Q-P} = (Q', P', E'_{Q-P})$  和节点集  $Q'_{11}, Q'_{21}, Q'_{31}, Q'_{41}, Q'_{12}, Q'_{22}, Q'_{32}, Q'_{42}$ , 原始互联网拓扑图  $G = (V, E)$ 。

[0350] 输出: 合并子图  $G'_m$ , 内核子图  $G'_{Core}$ 。

[0351] 符号定义:  $||S||$  表示集合  $S$  的势, 即集合  $S$  包含元素的总数。

[0352]  $d_v(G)$  表示节点  $v$  在图  $G$  的度。

[0353]  $N_v(G)$  表示节点  $v$  在图  $G$  的相邻节点集。

[0354]  $\lfloor x \rfloor$  表示元素  $0$  到  $x$  之间最接近于  $x$  的整数。

[0355]  $x \leftarrow y$  表示将右边的函数或数值  $y$  赋值给左边的变量  $x$ 。

[0356]  $\in$  表示“属于”;  $\exists$  表示“存在”;  $\cup$  表示集合的并集;  $\cap$  表示集合的交集。

[0357]  $A/B$ : 如果  $A$  和  $B$  为集合, 则表示从集合  $A$  中删除集合  $B$  的所有元素。

[0358]  $A \subseteq B$  表示集合  $A$  中的所有元素都被包含于集合  $B$ 。

[0359]  $\Phi$  表示不包含任何元素的空集。

[0360] 方法的流程与步骤:

[0361] S1. 从节点集  $Q'/Q'_{11}$  中均等概率地抽取  $||Q'_{12}||$  个节点  $Q_s = \{q_1, q_2, \dots\}$ , 并将采样二分图  $G'_{Q_1-II} = (II', Q', E'_{Q_1-II})$  节点集  $Q'_1$  包含的节点子集  $Q'_{12}$  替换为  $Q_s$ 。类似地, 采用从  $Q'/Q'_{21}, Q'/Q'_{31}$  和  $Q'/Q'_{41}$  抽取的节点集, 分别替换  $G'_{Q_2-BI} = (BI', Q_2', E'_{Q_2-BI})$  的节点子集  $Q'_{22}, G'_{Q_3-PI} = (PI', Q_3', E'_{Q_3-PI})$  的节点子集  $Q'_{32}$  和  $G'_{Q_4-RI} = (RI', Q_4', E'_{Q_4-RI})$  的节点子集  $Q'_{42}$ 。通过替换之后, 可以满足  $Q', Q'_2, Q'_3, Q'_4 \subseteq Q'$ 。

[0362] S2. 计算节点集  $V''_m = II' \cup BI' \cup PI' \cup RI' \cup QI' \cup Q' \cup P'$  和边集  $E''_m = E'_{Q_1-II} \cup E'_{Q_2-BI} \cup E'_{Q_3-PI} \cup E'_{Q_4-RI} \cup E'_{QI-PI} \cup E'_{Q-P}$ 。则, 合并图  $G''_m = (V''_m, E''_m)$  为所有采样二分图和采样匹配图的合并结果。

[0363] S3. 设  $V'_{Q, QI}$  为合并图  $G''_m$  中所有  $Q$  和  $QI$  节点组成的节点集 (注: 第1.1节“互联网拓扑图分解”已经将拓扑图中全部节点标记为  $P, Q, II, BI, QI, PI$  和  $RI$  七类不同的节点, 系统噪声  $OI$  节点已被从拓扑图中删除), 并设  $V'_{QI}$  为合并图  $G''_m$  中所有  $QI$  节点组成的节点集。定义  $G'_{Core} = (V'_{Q, QI}, E'_{Q, QI})$  为原始拓扑  $G$  中由节点集  $V'_{Q, QI}$  及这些节点之间的边生成的子图, 并定义  $G'_{inner} = (V'_{QI}, E'_{QI})$  为原始拓扑  $G$  中由节点集  $V'_{QI}$  及这些节点之间的边生成的子图。

[0364] S4. 计算节点集  $\bar{V}_Q = \{v \in V'_{Q, QI} \cap Q \mid d_v(G''_m) = 1\}$  和  $\bar{V}_{QI} = \{v \in V'_{QI} \mid d_v(G''_{inner}) = 1\}$ , 其中  $G''_{inner}$  为图  $G''_m$  中由所有  $PI, RI, QI$  节点及这些节点之间的边生成的子图。将节点集  $\bar{V}_Q$  分解为  $\bar{V}_Q^1 = \{v \in \bar{V}_Q \mid \exists w \in V'_{Q, QI}, (w, v) \in E'_{Q, QI}\}$  和  $\bar{V}_Q^2 = \bar{V}_Q / \bar{V}_Q^1$ , 并将节点集  $\bar{V}_{QI}$  分解为  $\bar{V}_{QI}^1 = \{v \in \bar{V}_{QI} \mid \exists w \in V'_{QI}, (w, v) \in E'_{QI}\}$  和  $\bar{V}_{QI}^2 = \bar{V}_{QI} / \bar{V}_{QI}^1$ 。

[0365] S5. 初始化边集  $\tilde{E}_{Core} = \Phi$  和  $\hat{E}_{Core} = \Phi$ 。对每个节点  $v \in \bar{V}_Q^1$ , 均等概率地选择一个节点  $w \in N_v(G'_{Core})$ , 并更新  $\tilde{E}_{Core} \leftarrow \tilde{E}_{Core} \cup \{(w, v)\}$ , 其中  $(w, v) \in E'_{Q, QI}$ 。此外, 对每个节点  $v \in \bar{V}_{QI}^1$ ,

均等概率地选择一个节点  $w \in N_v(G''_{inner})$ ，并更新  $\tilde{E}_{Core} \leftarrow \tilde{E}_{Core} \cup \{(w, v)\}$ ，其中  $(w, v) \in E'_{QI} \subseteq E'_{Q, QI}$ 。

[0366] S6. 均等概率地生成  $x$  条连接节点集  $\bar{V}_Q^2$  中  $2x$  个节点 (定义  $V_x$  为这  $2x$  个节点的集合) 的边  $E_Q^g = \{(w_1, v_1), (w_2, v_2), \dots\}$ ，其中  $x = \lfloor \|\bar{V}_Q^2\|/2 \rfloor$ ，并更新  $\hat{E}_{Core} \leftarrow \hat{E}_{Core} \cup E_Q^g$ 。如果  $\exists v, v \in \bar{V}_Q^2/V_x$ ，则均等概率地选择一个节点  $w \in V'_{Q, QI}/\{v\}$ ，并更新  $\hat{E}_{Core} \leftarrow \hat{E}_{Core} \cup \{(w, v)\}$ 。

[0367] S7. 均等概率地生成  $y$  条连接节点集  $\bar{V}_Q^2$  中  $2y$  个节点 (定义  $V_y$  为这  $2y$  个节点的集合) 的边  $E_Q^g = \{(w'_1, v'_1), (w'_2, v'_2), \dots\}$ ，其中  $y = \lfloor \|\bar{V}_Q^2\|/2 \rfloor$ ，并更新  $\hat{E}_{Core} \leftarrow \hat{E}_{Core} \cup E_Q^g$ 。如果  $\exists v, v \in \bar{V}_Q^2/V_y$ ，则均等概率地选择一个节点  $w \in V'_{QI}/\{v\}$ ，并更新  $\hat{E}_{Core} \leftarrow \hat{E}_{Core} \cup \{(w, v)\}$ 。

[0368] S8. 更新  $G'_{Core} = (V'_{Q, QI}, E'_{Q, QI})$ ： $E'_{Q, QI} \leftarrow E'_{Q, QI}/\hat{E}_{Core}$ ；并计算  $G'_m = G''_m \cup \tilde{E}_{Core} \cup \hat{E}_{Core}$ 。则  $G'_{Core}$  和  $G'_m$  为最终的输出结果。

[0369] 1.14 内核图采样方法

[0370] 输入：原始互联网拓扑  $G = (V, E)$ ，第1.13节输出的合并子图  $G'_m$  和内核子图  $G'_{Core} = (V'_{Core}, E'_{Core})$ 。

[0371] 输出：内核边集  $E''_{Core}$ 。

[0372] 符号定义： $\|S\|$  表示集合  $S$  的势，即集合  $S$  包含元素的总数。

[0373]  $d_v(G)$  表示节点  $v$  在图  $G$  的度。

[0374]  $\in$  表示“属于”； $\max(x, y)$  表示取  $x$  和  $y$  中的最大值。

[0375] 平均节点度是指图中全部节点的度之和除以全部节点的总数。

[0376] 方法的流程与步骤：

[0377] S1. 计算原始拓扑  $G$  的平均节点度  $d_{avg}$ ，计算合并子图  $G'_m$  的边数  $m(G'_m)$  和节点数  $n(G'_m)$ 。将所有边  $(w, v) \in E'_{Core}$  按照  $\max(d_w(G'_{Core}), d_v(G'_{Core}))$  不降顺序排列。

[0378] S2. 从边集  $E'_{Core}$  中删除排在前面的  $\|E'_{Core}\| - d_{avg} \cdot n(G'_m)/2 + m(G'_m)$  条边，并将剩余边组成的集合定义为  $E''_{Core}$  (其包含  $d_{avg} \cdot n(G'_m)/2 - m(G'_m)$  条边)。

[0379] S3.  $E''_{Core}$  为最终输出的内核边集。

[0380] 1.15 互联网拓扑最终采样结果

[0381] 输入：第1.13节输出的合并子图  $G'_m$ ，第1.14节输出的内核边集  $E''_{Core}$ 。

[0382] 输出：最终采样结果  $G'$ 。

[0383] 符号定义： $\cup$  表示集合的并集。

[0384] 方法的流程与步骤：

[0385] S1. 真实世界互联网拓扑的最终采样结果为  $G' = G'_m \cup E''_{Core}$ 。

[0386] 实施例以互联网测试床的组播路由协议测试任务为案例。选择该任务需求的原因在于组播路由协议测试指标对测试床拓扑结构的变化比较敏感。实施例以Stanford网络数据池 (<http://snap.stanford.edu/data/>) 和ITDK工程 (<http://www.caida.org/>) 提供的三

个真实世界自治系统级互联网拓扑的探测数据为原始拓扑。这三个原始拓扑分别来源于AS-733(2000年01月探测)、AS-Caida(2007年11月探测)和ITDK(2017年02月探测)。

[0387] 测试任务首先采用本发明的拓扑采样技术,对真实世界探测得到的原始拓扑进行规模缩减(要求在保持节点度分布、聚类系数、RichClub连通性和路径长度等重要拓扑属性近似稳定不变前提下,实现大于96%的规模缩减率),得到小规模拓扑结构的互联网测试床;然后在小规模测试床上计算组播路由协议测试指标(包含延时率和费用率)的测量数值;最后通过与真实世界拓扑结构上测试结果的对比,分析本发明(拓扑规模缩减技术)的实用性,即是否能够由小规模测试床上的测试结论等效推演得到组播路由协议在大规模真实网络的运行效果。

[0388] 1、组播路由协议测试

[0389] 互联网组播架构可以划分为组播路由协议、组播传输协议和群组管理协议三类。其中组播路由协议构建组播拓扑(即,群组内源节点到多个接收节点的组播数据路径分布),组播传输协议管控组播拓扑上数据流的传输机制,组播管理协议负责群组成员的管理。因此,组播路由协议紧密依赖于互联网测试床的拓扑结构,组播拓扑通常是测试床拓扑的一个子集。目前,组播路由协议可以划分为基于源的树SBT(Source Based Tree)协议和基于中心的树CBT(Center Based Tree)协议两类。SBT和CBT采用不同的机制生成包含群组全部成员的组播拓扑,其中SBT由许多最短路径树的并集组成(每个树以群组内的一个数据发送源节点为根,形成了从该源节点发送数据的组播路径);反之,CBT由唯一的最短路径树组成,该树被群组内所有的成员节点共享。

[0390] 以某个源节点为根的最短路径树,由该源节点至群组中所有接收节点的最短路径组成。一般地,从一个源节点到一个接收节点的延时,采用它们之间的最短路径长度来测量。因此,SBT可以得到从每个源节点至所有接收节点的最小平均延时。然而,费用是评估该协议的另一个重要指标,其被定义为组播分布树(即,组播拓扑)覆盖互联网测试床拓扑的链路(边)总数。易知,CBT可以得到较优的费用指标。虽然如何使平均延时和费用之间达到平衡仍是组播路由领域的一个开放问题,但是CBT选择网络拓扑的“中心”作为其唯一最短路径树的根节点,已成为当前应用较为广泛的组播路由协议。以什么标准定义“网络中心”,以及“网络中心”如何选择,仍是CBT协议簇的一个重要研究方向。

[0391] 目前,CBT协议的评价通常以SBT协议为基准,因为SBT可以达到最优(小)的平均延时指标。常用的CBT协议评价指标包括:

[0392] 延时率:被定义为SBT的平均延时与CBT的平均延时的比率。

[0393] 费用率:被定义为SBT的费用与CBT的费用的比率。

[0394] 选择“网络中心”(即,唯一最短路径树的根节点)对于CBT协议至关重要。然而,本实施例并不关心“网络中心”的选择方法。精确地计算“网络中心”已被证明是NP完全问题;也就是说,这种精确计算难以被用于大规模网络的协议设计。因此,CBT协议的网络中心通常采用启发式的方法计算。

[0395] 本实施例重点关心自治系统级互联网测试床拓扑的规模缩减,不关心被试CBT技术的优化。因此,仅考虑一个具体的CBT协议,其“网络中心”的选择策略为:标准分布地随机选择一个源节点作为“网络中心”。标准分布是指分布概率值均等的概率分布。

[0396] 2、测试任务验证结果

[0397] 拓扑属性验证结果:

[0398] 本发明的技术针对AS-733(2000年01月探测,包含6,474节点)、AS-Caida(2007年11月探测,包含26,389节点)和ITDK(2017年02月探测,包含41,626节点)三个真实世界的自治系统级互联网拓扑,都可以在保证节点度分布、聚类系数、RichClub连通性和路径长度分布等重要拓扑属性近似稳定不变的前提下,实现96%以上的节点规模缩减。

[0399] 组播路由协议测试验证结果:

[0400] 针对上述AS-733、AS-Caida和ITDK三个真实世界原始拓扑,依据本发明技术采样得到的小规模拓扑结构(节点规模缩减率大于96%),在互联网测试床上配置拓扑连接关系。通过实验对比验证得到:在小规模测试床上组播路由协议的测试结论与在大规模真实网络上组播路由协议的运行效果基本保持一致。

[0401] 本公开的实施例提供的一种互联网测试床拓扑结构大比例规模缩减装置,该实施例的一种互联网测试床拓扑结构大比例规模缩减装置包括:处理器、存储器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,例如拓扑程序。所述处理器执行所述计算机程序时实现上述各个方法实施例中的步骤。

[0402] 所述装置包括:存储器、处理器以及存储在所述存储器中并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序的装置包括:

[0403] 建模单元,用于获取真实世界互联网拓扑结构的探测数据,以简单无向图 $G=(V, E)$ 建模自治系统级互联网拓扑结构,其中集合 $V$ 中节点的物理意义为真实世界互联网的自治系统域,集合 $E$ 中边的物理意义为不同自治系统域之间的数据通信路径;

[0404] 分解拓扑单元,用于采用互联网拓扑结构模型分解互联网拓扑图 $G=(V, E)$ 为七个二分图 $Q_1-II$ 、 $Q_2-BI$ 、 $Q_{I_1}-RI$ 、 $Q_4-RI_1$ 、 $QI-PI$ 、 $Q_3-PI$ 、 $Q-P$ ,一个匹配图 $BI-BI$ 和一个内核图Core;

[0405] 二分图采样1单元,用于调用二分图采样方法1对二分图 $Q_1-II$ 、二分图 $Q_2-BI$ 和二分图 $Q_{I_1}-RI$ 进行采样,从而抽取部分节点和边构成采样子图;

[0406] 匹配图采样单元,用于调用匹配图采样方法对匹配图 $BI-BI$ 进行采样,抽取部分节点和边构成采样子图;

[0407] 二分图采样2单元,用于调用二分图采样方法2对二分图 $Q_4-RI_1$ 进行采样,抽取部分节点和边构成采样子图;

[0408] 二分图采样3单元,用于调用二分图采样方法3对二分图 $QI-PI$ 和二分图 $Q-P$ 进行采样,抽取部分节点和边构成采样子图;

[0409] 二分图采样4单元,用于调用二分图采样方法4对二分图 $Q_3-PI$ 进行采样,抽取部分节点和边构成采样子图;

[0410] 子图合并单元,用于对得到的所有采样子图进行合并,得到互联网外围拓扑所有采样子图的合并图;

[0411] 内核图采样单元,用于调用内核图采样方法对内核图进行采样,从而抽取部分节点和边构成采样子图;

[0412] 采样子图获取单元,用于对得到的外围拓扑采样子图的合并图和内核拓扑的采样子图进行合并,从而得到真实世界互联网拓扑图的采样子图;

[0413] 仿真测试单元,依据得到的真实世界互联网拓扑图的规模缩减采样子图,配置互

联网测试床的拓扑连接关系,从而调度测试床有限的仿真节点资源,生成节点规模大比例缩减的互联网仿真测试环境。

[0414] 所述一种互联网测试床拓扑结构大比例规模缩减装置可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述一种互联网测试床拓扑结构大比例规模缩减装置可包括,但不仅限于,处理器、存储器。本领域技术人员可以理解,所述例子仅仅是一种互联网测试床拓扑结构大比例规模缩减装置的示例,并不构成对一种互联网测试床拓扑结构大比例规模缩减装置的限定,可以包括比例更多或更少的部件,或者组合某些部件,或者不同的部件,例如所述一种互联网测试床拓扑结构大比例规模缩减装置还可以包括输入输出设备、网络接入设备、总线等。

[0415] 所称处理器可以是中央处理单元(Central Processing Unit,CPU),还可以是其他通用处理器、数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现成可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等,所述处理器是所述一种互联网测试床大比例规模缩减装置的控制中心,利用各种接口和线路连接整个一种互联网测试床拓扑结构大比例规模缩减装置的各个部分。

[0416] 所述存储器可用于存储所述计算机程序和/或模块,所述处理器通过运行或执行存储在所述存储器内的计算机程序和/或模块,以及调用存储在存储器内的数据,实现所述一种互联网测试床拓扑结构大比例规模缩减装置的各种功能。所述存储器可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据手机的使用所创建的数据(比如音频数据、电话本等)等。此外,存储器可以包括高速随机存取存储器,还可以包括非易失性存储器,例如硬盘、内存、插接式硬盘,智能存储卡(Smart Media Card,SMC),安全数字(Secure Digital,SD)卡,闪存卡(Flash Card)、至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

[0417] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0418] 尽管本公开的描述已经相当详尽且特别对几个所述实施例进行了描述,但其并非旨在局限于任何这些细节或实施例或任何特殊实施例,而是应当将其视作是通过参考所附权利要求考虑到现有技术为这些权利要求提供广义的可能性解释,从而有效地涵盖本公开的预定范围。此外,上文以发明人可预见的实施例对本公开进行描述,其目的是为了提供有用的描述,而那些目前尚未预见的对本公开的非实质性改动仍可代表本公开的等效改动。

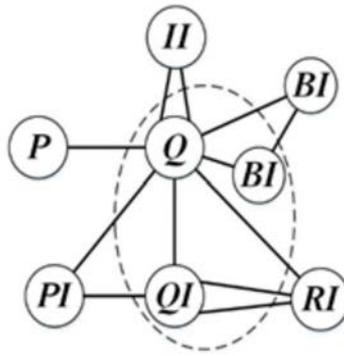


图1

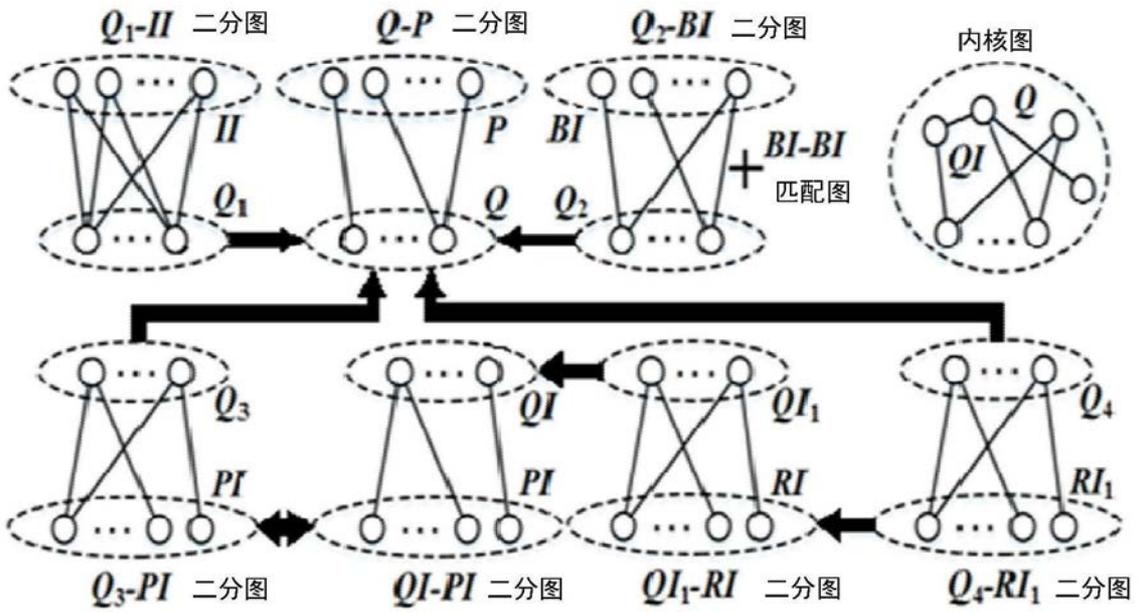


图2

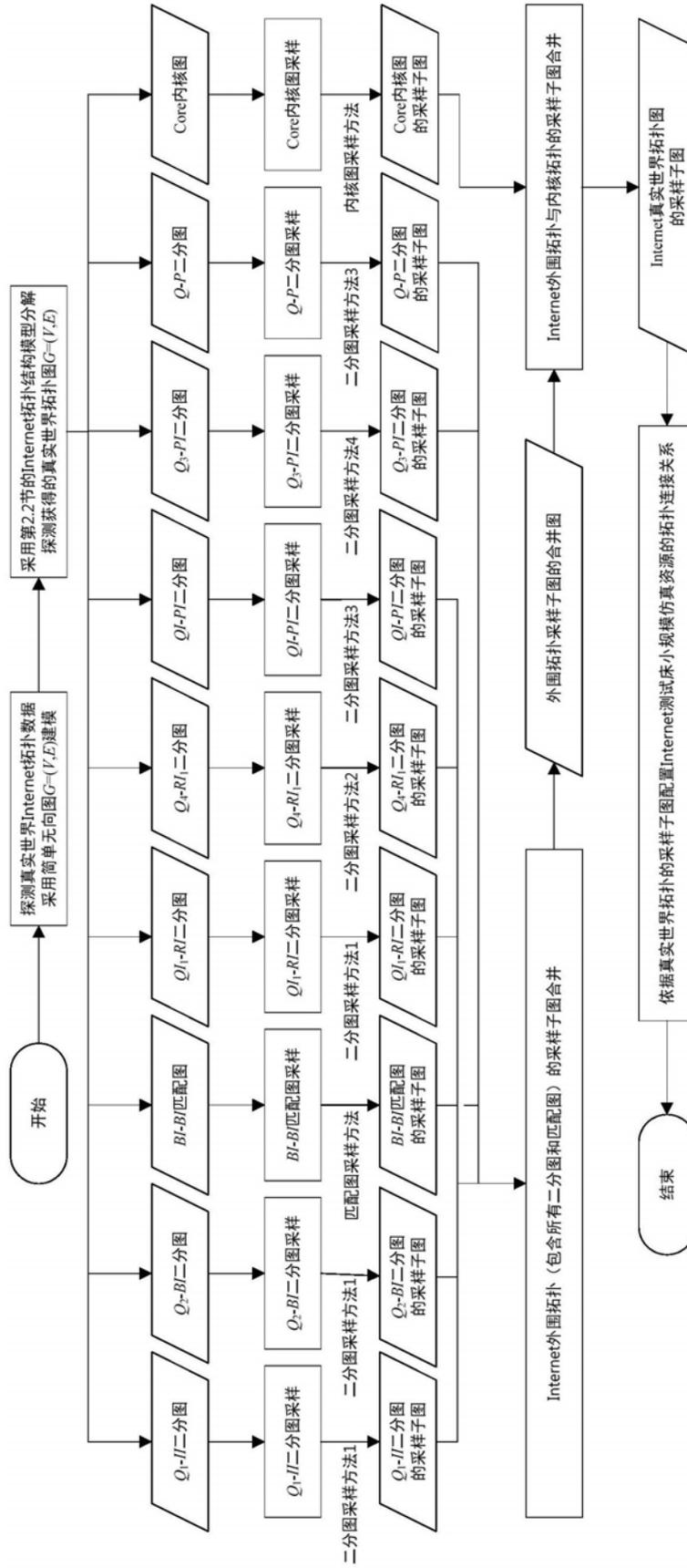


图3

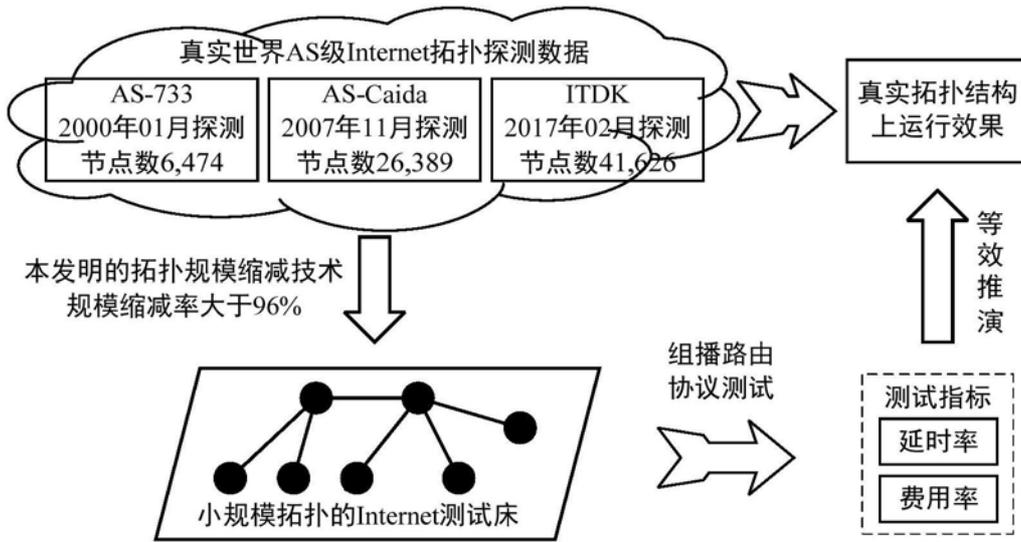


图4