

基于遗传算法的层间等级网络编码优化^①司菁菁^② 孙明明 程银波

(燕山大学信息科学与工程学院 秦皇岛 066004)

摘 要 面向单信源异构信宿网络,研究了层间等级网络编码的编码类型优化。基于遗传算法,提出了一种最优编码类型的快速搜索方案。该方案充分考虑了信源输出链路上进行的层间等级网络编码的编码类型对整个网络传输性能的影响,将网络总吞吐量作为评价编码类型优劣的标准,设计了符合层间等级网络编码本质特性的遗传操作。实验结果表明,与分层组播网络编码和基于现有启发式算法的层间等级网络编码相比,基于本文方案实现的层间等级网络编码能够为单信源异构信宿网络获得更高的网络总吞吐量。

关键词 网络编码,多速率,组播网络编码,层间等级网络编码,遗传算法

0 引言

近年来,学者们结合信源分层编码技术,提出了多种多速率网络编码方案,用以实现单信源异构信宿网络中的高效信息传输。分层组播(layered multicast)^[1-4]是一类基于组播内(intra-session)网络编码的多速率组播方案。此类方案仅允许对属于同一信源编码层的数据进行网络编码,因此,需要在网络中为每个信源编码层的组播划分出独立的编码子图。然而,这一条件在一般的网络中未必能够满足。层间等级网络编码(inter-layer hierarchical network coding)^[5-7]是一种特殊的组播间(inter-session)多速率网络编码方案^[8-10],这种方案允许不同信源编码层的数据,按照重要性由高到低的顺序共同参与网络编码。与分层组播相比,层间等级网络编码能够获得更高的网络资源利用率。

在异构信宿网络中利用层间等级网络编码实现多速率数据传输的关键问题是如何为网络中的各链路确定其上传输数据的最佳编码类型,从而尽可能充分地利用各信宿的接收能力,实现高质量的信源数据重建。文献[6]提出了一种层间等级网络编码类型优化的启发式算法,它给出了网络中的每个信宿寻找传输各类编码数据的独立路径。当为不同信

宿传输不同类型编码数据的路径具有公共链路时,该算法可为此链路选取最低的可行编码类型。因此,该算法会以牺牲高带宽信宿的吞吐量为代价来提高低带宽信宿的吞吐量,对于一般网络,未必能够达到网络的最大总吞吐量。针对上述问题,本文面向单信源异构信宿网络,研究了层间等级网络编码的编码类型优化。基于遗传算法,提出了一种最优编码类型的快速搜索方案。该方案充分考虑了信源输出链路上的编码类型对整个网络传输性能的决定性影响,针对层间等级网络编码的本质特性设计了具体的遗传操作。实验结果表明,与分层组播^[1]和基于现有启发式算法的层间等级网络编码^[6]相比,基于本文方案实现的层间等级网络编码,能够使单信源异构信宿网络获得更高的网络总吞吐量。

1 问题模型

本文讨论单信源有向无环网络 $G = (V, E)$, 其中 V 表示节点集, E 表示链路集。令 $V = \{s\} \cup M \cup T$, 其中 s 为信源节点, M 为中间节点集, T 为信宿节点集。设任意链路 $e \in E$ 均具有单位容量。令 $In(v)$ 和 $Out(v)$ 分别表示节点 $v \in V$ 的输入链路集和输出链路集。若链路 $e_1 \in Out(v_1)$, $e_\tau \in In(v_{\tau+1})$, 且对于任意 $2 \leq i \leq \tau$ 均存在节点 $v_i \in V$

① 国家自然科学基金(61303128), 河北省自然科学基金(F2014203183), 河北省高等学校科学技术研究项目(Q2012087), 燕山大学青年教师自主研究计划课题(13LGB015)资助项目。

② 女, 1980 年生, 博士, 副教授; 研究方向: 多媒体通信, 网络编码等; 联系人, E-mail: sijingjing@gmail.com (收稿日期: 2013-02-21)

使得 $e_{i-1} \in In(v_i)$ 、 $e_i \in Out(v_i)$, 则称链路序列 $e_1, e_2, \dots, e_\tau$ 为一条由节点 v_1 到节点 $v_{\tau+1}$ 的长为 τ 的路径。若两条路径没有共同的链路, 则称这两条路径是非重叠的。

设信源节点 s 利用分层编码算法将信源数据编码成 L 层。第 1 层为基本层, 具有最高的重要性; 第 2 层到第 L 层为增强层, 重要性逐层降低。根据信源分层编码的特性, 某信宿接收到的信源层数越多, 其重建的信源数据的质量越高。令 F_q 表示阶为 q 的有限域。设信源节点单位时间内生成的第 k 层信源编码数据可以表示成有限域 F_q 上的 r_k 个数据 $x_{k,1}, x_{k,2}, \dots, x_{k,r_k}$, 并将 r_k 称为第 k 信源层的速率。令

$$A_k = \sum_{i=1}^k r_i, \quad k = 1, 2, \dots, L \quad (1)$$

分层组播仅允许对属于同一信源编码层的数据进行网络编码。一个第 k 层的分层组播网络编码数据可以表示为

$$\mathbf{y}'_k = [x_{k,1}, x_{k,2}, \dots, x_{k,r_k}] \cdot \mathbf{f}'_k \quad (2)$$

其中, r_k 维编码系数向量

$$\mathbf{f}'_k = [f'_{k,1}, f'_{k,2}, \dots, f'_{k,r_k}]^T \in F_q^{r_k} \quad (3)$$

可见, 在第 k 层的分层组播网络编码数据中, 仅包含了第 k 信源层的信息。

为了进一步提高异构信宿网络中的资源利用率, 本文研究了层间等级网络编码, 将第 k 类层间等级网络编码定义为前 k 层信源编码数据的线性组合。因此, 一个第 k 类层间等级网络编码数据 $\mathbf{y}_k$ 可以表示为

$$\mathbf{y}_k = [x_{1,1}, \dots, x_{1,r_1}, x_{2,1}, \dots, x_{2,r_2}, \dots, x_{k,1}, \dots, x_{k,r_k}] \cdot \mathbf{f}_k \quad (4)$$

其中, A_k 维编码系数向量

$$\mathbf{f}_k = [f_{k,1}, f_{k,2}, \dots, f_{k,A_k}]^T \in F_q^{A_k} \quad (5)$$

可见, 与分层组播不同, 在第 k 类层间等级网络编码数据中, 包含了前 k 个信源层的信息。若某信宿 $t \in T$ 能够接收到 A_k 个前 k 类层间等级网络编码数据, 且这些编码数据对应的 A_k 维编码系数向量是线性无关的, 则信宿 t 能够解码重建出前 k 个信源层。当确定编码系数向量的线性无关性时, 对于 $1 \leq i < k$ 的第 i 类编码数据, 需要将其对应的 A_i 维编码系数向量补零加长成 A_k 维。

本文面向单信源异构信宿网络, 研究了层间等级网络编码的编码类型优化。基于遗传算法, 为网络中的各链路, 搜索其上传送的层间等级网络编码数据的最佳类型。然后, 利用随机网络编码, 实现各

链路指定类型的层间等级网络编码。当为链路 $e \in E$ 确定的最佳编码类型为 k 时, e 上传送的数据为

$$[x_{1,1}, \dots, x_{1,r_1}, x_{2,1}, \dots, x_{2,r_2}, \dots, x_{k,1}, \dots, x_{k,r_k}] \cdot \mathbf{f}_k$$

其中, A_k 维编码系数向量 $\mathbf{f}_k$ 是在有限域 F_q 上随机生成的。

为了便于优化问题的描述, 进行如下变量的定义。令 $z_{t,k}$ 表示信宿 $t \in T$ 是否能够解码出第 k 信源编码层, 即

$$z_{t,k} = \begin{cases} 1, & \text{若信宿 } t \text{ 能够解码出第 } k \text{ 信源层} \\ 0, & \text{若信宿 } t \text{ 不能解码出第 } k \text{ 信源层} \end{cases} \quad (6)$$

按照信源分层编码的特点, 此变量应满足 $z_{t,k} \geq z_{t,k+1}$, 即信宿 t 只有在保证前 k 层信源数据可解的情况下, 才能解码出第 $k+1$ 层信源数据。

令 R_t 表示信宿 t 的吞吐量。根据 t 对各信源编码层的解码情况, R_t 可以表示为

$$R_t = \sum_{k=1}^L r_k \cdot z_{t,k} \quad (7)$$

网络中全部信宿的总吞吐量可以表示为

$$\sum_{t \in T} R_t = \sum_{t \in T} \sum_{k=1}^L r_k \cdot z_{t,k} \quad (8)$$

在下文中, 若某链路上传输的层间等级网络编码数据的类型为 k , 则称此链路的编码类型为 k 。令 $Type_e$ 表示链路 e 的编码类型, 则 $Type_e$ 的可能取值为 $1, 2, \dots, L$ 。

2 基于遗传算法的编码类型优化

本文基于遗传算法, 研究层间等级网络编码的编码类型优化分配, 设计了一种最优编码类型的快速搜索方案。本方案以网络中各链路上进行的层间等级网络编码的编码类型作为优化对象, 以网络中全部信宿总吞吐量 $\sum_{t \in T} R_t$ 的最大化作为优化目标。方案流程如图 1 所示。

2.1 预处理

本方案利用染色体中的一个基因表示网络中一条链路的编码类型。实验中发现, 当随机设置基因的排列顺序, 即随机选择链路的排列顺序时, 遗传算法的收敛速度较低, 且搜索到的最优解的质量不高。更重要的是, 信源节点输出链路 $e \in Out(s)$ 的编码类型, 对整个网络的传输性能影响非常大。因此, 本文按照网络中链路的拓扑顺序设置染色体: 将信源节点输出链路对应的基因, 排在染色体的最前面; 对

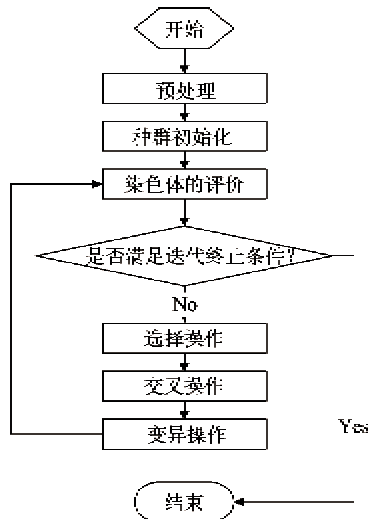


图1 方案流程图

于网络中的任意节点 $v \in V$, 将各输入链路 $d \in In(v)$ 对应的基因排在所有输出链路 $e \in Out(v)$ 对应的基因之前。这样既能够保证遗传算法的效率, 又便于在遗传过程中对染色体中的部分基因进行调整, 保证染色体的合理性。在本方案中, 合理的染色体应满足如下条件: 对于任意节点 $v \in V$, 任意输出链路 $e \in Out(v)$ 的编码类型 $Type_e$, 均小于等于任意输入链路 $d \in In(v)$ 的编码类型 $Type_d$ 。

此外, 预处理过程还包括对种群规模、交叉概率、变异概率、迭代终止参数等的设置。

2.2 种群初始化

本方案要求初始化种群中的所有染色体均满足合理性条件。具体设置方法如下:

(1) 按照链路的拓扑顺序, 设置染色体基因。令信源输出链路集 $Out(s)$ 中的链路对应染色体中的前 $|Out(s)|$ 个基因。

(2) 在整数区间 $[1, L]$ 的范围内, 为前 $|Out(s)|$ 个基因随机赋值。

(3) 为染色体中的其他 $|E| - |Out(s)|$ 个基因赋值: 对于某中间节点 v 的输出链路 $e \in Out(v)$, 判定各输入链路 $d \in In(v)$ 的编码类型 $Type_d$ 是否已经全部确定, 若 $\{Type_d: d \in In(v)\}$ 已全部确定, 则从其中随机选择一个, 为 $Type_e$ 赋值。此过程重复进行, 直到为所有基因赋值完毕为止。

利用上述过程, 生成 $2P_{size}$ 条初始染色体, 然后将它们按照适应度函数值的大小降序排列, 选择前 P_{size} 条染色体形成初始种群。

2.3 适应度函数

遗传算法是利用适应度函数来评价染色体的优

劣的。本文的优化目标是使网络中的异构信宿获得最大的总吞吐量。因此, 将染色体 φ 的适应度函数定义为

$$f(\varphi) = \sum_{t \in T} R_t(\varphi) \quad (9)$$

式中 $R_t(\varphi)$ 表示染色体 φ 代表的层间等级网络编码类型分配状态为信宿 t 获得的吞吐量, $f(\varphi)$ 表示 φ 为所有信宿获得的总吞吐量。因此, $f(\varphi)$ 值越大, 染色体 φ 越接近于最优解。

根据染色体 φ 计算 $R_t(\varphi)$ 的流程如图 2 所示。在图 2 中, $p_{t,k}(\varphi)$ 表示在染色体 φ 对应的编码类型分配状态中, 为信宿 t 传输第 k 类编码数据的非重叠路径的条数。设链路 $e_1, e_2, \dots, e_r$ 形成了一条由信源 s 到信宿 t 的长为 r 的路径。若染色体 φ 为链路 $e_1, e_2, \dots, e_r$ 分配的编码类型均为 k , 则此路径为信宿 t 传输的编码数据的类型为 k 。

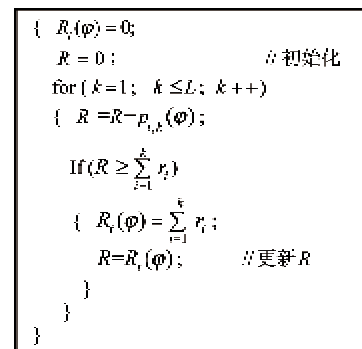


图2 计算 $R_t(\varphi)$ 的流程

2.4 遗传操作

选择。结合最优保持策略与轮盘赌原则, 为下一代种群选择合适的染色体。首先, 在当前种群中, 选择适应度值最大的 n_{opt} 条染色体直接进入下一代; 然后, 根据轮盘赌原则, 以 $f(\varphi_i) / \sum_{i=1}^{P_{size}} f(\varphi_i)$ 为概率, 选择 $P_{size} - n_{opt}$ 条染色体进入下一代。在新一代种群中, 染色体的数量仍为 P_{size} 。

交叉。采用两点交叉法, 并将交叉操作限定在那些对应于信源节点输出链路的基因上。为了保证染色体的合理性, 对交叉生成的染色体进行验证: 与信源节点输出链路对应的前 $|Out(s)|$ 个基因保持不变; 依次判断后续各基因是否合理, 直至全部基因均合理为止。若某链路 $e \in Out(v)$ 对应的基因大于链路集 $\{d: d \in In(v)\}$ 中各链路对应的基因, 则 e 对应的基因不合理。调整方法: 从 $\{d: d \in In(v)\}$ 中各链路对应的基因中随机选择一个, 将其值赋给

e 对应的基因。

变异。根据染色体的特性,采用两种变异操作:基因逆序和单个基因突变,并将变异限定在前 $|Out(s)|$ 个基因范围内。为保证变异后染色体的合理性,采用与交叉操作后相同的处理方法,遍历染色体中的所有基因,进行局部调整。

2.5 遗传终止条件

采用如下两个遗传终止条件:(1)迭代次数达到预处理过程中设置的最大遗传代数 N_g ; (2)连续 N_{stop} 代种群中最优染色体的适应度函数值保持不变。只要满足上述条件中的一个,遗传搜索过程结束。当前种群中的最优染色体即为搜索结果。最优染色体对应的层间等级网络编码类型分配状态,即为本次搜索获得的最优编码类型分配。

3 实验结果

以随机生成的单信源有向无环网络作为实验对象,利用C++语言编写实验程序,测试方案性能。首先,利用本文方案搜索网络中各链路上进行层间等级网络编码的最优编码类型;然后,采用随机线性网络编码,为网络中的各链路实现指定类型的层间等级网络编码;最后,分析本方案获得的网络总吞吐量,并与现有算法进行比较。

3.1 实验1

实验网络的拓扑结构如图3所示。每条链路具有单位容量。信宿 t_1, t_2, t_3 的最大流分别是1、4、1。设信源节点 s 利用分层编码算法将信源数据编码成2层,码率分别为 $r_1 = 1, r_2 = 3$ 。

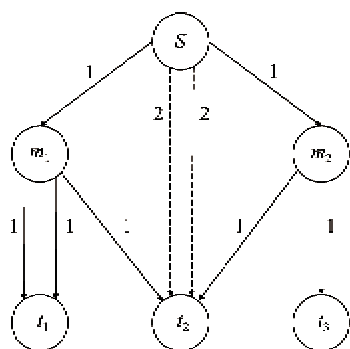


图3 实验1的网络拓扑结构以及采用文献[1]的算法获得的分层组播方案

当采用文献[1]的分层组播方案对各信源编码层进行独立组播时,最佳传输方案如图3所示。其中,实线表示传输第1层网络编码数据的链路,虚线

表示传输第2层网络编码数据的链路。可见,信宿 t_1, t_2, t_3 均能解码出第1信源层,但均不能解码出第2信源层。网络的总吞吐量为3。

当采用文献[6]提出的基于启发式算法的层间等级网络编码时,获得的最佳编码类型分配如图4所示。图中,每条链路旁边的整数表示此链路上传输的层间等级网络编码数据的类型。可见,信宿 t_1 接收到了1个第1类编码数据,可以解码出第1信源层;信宿 t_2 接收到了2个第1类编码数据、2个第2类编码数据,但仅能解码出第1信源层;信宿 t_3 接收到了1个第1类编码数据,可以解码出第1信源层。因此,网络的总吞吐量为3。

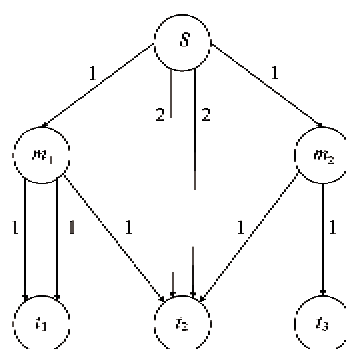


图4 文献[6]启发式算法获得的最优编码类型分配

当采用本文提出的基于遗传算法的层间等级网络编码类型分配方案时,获得的最佳编码类型分配如图5所示。图中,每条链路旁边的整数表示此链路上传输的层间等级网络编码数据的类型。可见,信宿 t_1 接收到了1个第1类编码数据,可以解码出第1信源层;信宿 t_2 接收到了1个第1类编码数据、3个第2类编码数据,可以解码出第1、2信源层;信宿 t_3 接收到了1个第2类编码数据,不能解码出有效的信源信息。因此,网络总吞吐量为5,高于前两方案的结果。

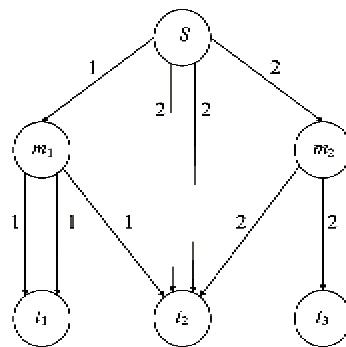


图5 本文算法获得的最优编码类型分配

3.2 实验 2

实验网络的拓扑结构如图 6 所示。各链路按如图所示的容量,转化成并行的单位容量链路。信宿 t_1 、 t_2 、 t_3 的最大流分别是 5、9、3。设 s 利用信源分层编码算法将信源信号编码成 3 层,码率分别为 $r_1 = 3$ 、 $r_2 = 2$ 、 $r_3 = 4$ 。

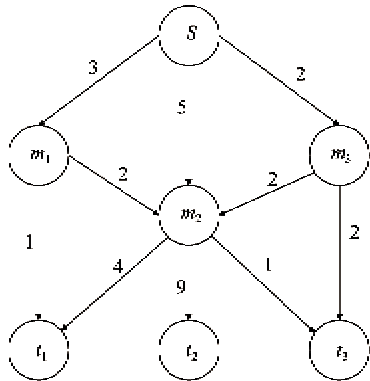


图 6 实验 2 的网络拓扑结构

分别采用文献[1]提出的分层组播网络编码方案、文献[6]提出的基于启发式算法的层间等级网络编码方案和本文提出的基于遗传算法搜索最优编码类型的层间等级网络编码方案,为图 6 所示的实验网络实现多速率数据传输。各方案获得的网络总吞吐量如表 1 所示。从表中可见,本文方案获得的网络总吞吐量高于前两方案的结果。

表 1 实验 2 的结果

方案名称	吞吐量
文献[1]的分层组播网络编码	13
文献[6]的层间等级网络编码	14
本文方案	17

3.3 对随机网络进行实验

随机构建单信源有向无环网络,并随机设置网络中各链路的容量。对于每个实验网络,首先,采用文献[2]的方案,确定信源分层编码的最优速率;然后,分别采用文献[1]的分层组播网络编码方案、文献[6]的层间等级网络编码方案和本文设计的层间等级网络编码方案实现信源分层编码数据的多速率传送;最后,比较这三种方案各自获得的网络总吞吐量。这里以 8 个随机网络为例,在表 2 中列出了实验结果。

表 2 针对 8 个随机网络的性能对比

网络序号	网络信宿最大流之和	文献[1]方案获得的总吞吐量	文献[6]方案获得的总吞吐量	本文方案获得的总吞吐量
1	7	6	6	7
2	11	8	8	11
3	18	14	14	18
4	28	18	22	28
5	33	21	25	33
6	36	24	32	36
7	43	30	37	43
8	48	36	40	48

实验结果表明,与分层组播和基于现有启发式算法的层间等级网络编码相比,本文提出的基于遗传算法实现编码类型优化分配的层间等级网络编码方案,能够有效提高异构信宿网络的总吞吐量。

4 结论

本文面向单信源异构信宿网络研究了层间等级网络编码的编码类型优化分配,提出了一种基于遗传算法的最优编码类型快速搜索方案,它体现了信源输出链路上的编码类型对整个网络传输性能的决定性影响。本研究以网络总吞吐量作为编码类型优劣的评价指标,针对层间等级网络编码的特性设计了具体遗传操作。实验结果表明,与分层组播和基于现有启发式算法的层间等级网络编码方案相比,基于本文方案实现的层间等级网络编码能够获得更高的网络总吞吐量。

本文方案为多速率线性网络编码的优化研究以及层间等级网络编码在异构信宿网络中的实际应用提供了一种思路。在实验中还发现,信源分层编码的速率对层间等级网络编码效率的影响也很大。如何利用进化算法,实现信源分层编码速率与层间等级网络编码类型分配的联合优化,是需要进一步研究的内容。

参考文献

[1] Zhao J, Yan F, Zhang Q, et al. LION: layered overlay multicast with network coding. *IEEE Trans on Multimedia*, 2006, 8(5): 1021-1032

[2] Si J, Zhuang B, Cai A. Optimization of layered multicast with network coding and genetic algorithm. *The Journal of China Universities of Posts and Telecommunications*, 2009, 16(3): 52-58

- [3] 林晓斌, 许胤龙, 詹成等. 基于网络编码的分层媒体多播中的层速率分配优化. 电子与信息学报, 2010, 32(10): 2421-2428
- [4] Gopinathan A, Li Z. Optimal layered multicast. *ACM Trans on Multimedia Computing, Communications and Applications*, 2011, 7(2): 1-24
- [5] Wu Y. Distributing layered content using network coding. In: Proceedings of the 5th IEEE Annual Communications Society Conference on Sensor, Mesh and Ad hoc Communications and Networks Workshops, San Francisco, USA, 2008
- [6] Si J, Zhuang B, Cai A. Hierarchical multicast with inter-layer random network coding. *High Technology Letters*, 2011, 17(1): 86-90
- [7] Shao M, Dumitrescu S, Wu X. Layered multicast with inter-layer network coding for multimedia streaming. *IEEE Trans on Multimedia*, 2011, 13(2): 353-365
- [8] Cogill R, Shrader B, Ephremides A. Stable throughput for multicast with inter-session network coding. In: Proceedings of MILCOM 2008, San Diego, CA, USA, 2008
- [9] Yang M, Yang Y. A linear inter-session network coding scheme for multicast. *Journal of Communications*, 2009, 11(4): 865-872
- [10] Widemer J, Capalbo A, Anta F, et al. Rate allocation for layered multicast streaming with inter-layer network coding. In: Proceedings of INFOCOM 2012, Orlando, Florida, USA, 2012. 2796-2800

Optimizing inter-layer hierarchical network codes based on the genetic algorithm

Si Jingjing, Sun Mingming, Cheng Yinbo

(School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004)

Abstract

The coding type optimization of the inter-layer hierarchical network coding for single-source heterogeneous-receivers networks was investigated. Based on the genetic algorithm, a strategy for fast searching the optimum coding type was proposed. This strategy emphasizes the optimization of the coding types on the outgoing links of the source node, and takes the maximization of the overall network throughput as the optimization objective. Special genetic processes were designed according to the characteristics of inter-layer hierarchical network codes. The experimental results show that compared to the layered multicast network coding and the inter-layer hierarchical network coding using the existing heuristic method, the inter-layer hierarchical network coding using the proposed optimization strategy can achieve the higher overall network throughputs for single-source heterogeneous-receiver networks.

Key words: network coding, multi-rate, inter-session network code, inter-layer hierarchical network code, genetic algorithm