

适用于线性网络编码关键路径的实时性算法

韩晓冬^{1,2} 高飞² 张立炜¹

1 中华人民共和国科学技术部信息中心 北京 100862

2 北京理工大学信息与电子学院 北京 100081

(hanxd@most.cn)

摘要 如今,人类社会存储和交换的信息总量呈几何级数飞速增长,数据传输的吞吐量和实时性亟待提升。然而,现有的网络编码研究专注于提升吞吐量,忽略了实时性对大数据网络多路径传输性能的重大影响。为此,文中针对线性网络编码的最快到达问题,提出一种矩阵优化相乘的关键路径算法,以提高算法的实时性。具体地,使用抽象代数分析关键路径算法,构造了关键路径的交换环代数,并证明了最优子结构性。仿真结果显示,随着网络节点个数 n 的增加,基于 Strassen 思想优化的关键路径算法能够极大地降低计算复杂度,成功将时间复杂度降至 $O(n^{2.81} \lg n)$,缩短了传播时延,提高了数据传输的实时性。当 $n > 6$ 时,相比基于重复平方关键路径算法,基于 Strassen 关键路径算法的时间开销的增长速率明显更低;特别地,当 $n = 12$ 时,基于 Strassen 关键路径算法的计算量约是基于重复平方关键路径算法的 $2/3$,而所需的时间开销约为后者的 $1/2$ 。

关键词: 关键路径; 代数结构; 交换环; 矩阵乘法; 线性网络编码

中图分类号 TP301.6

Novel Real-time Algorithm for Critical Path of Linear Network Coding

HAN Xiao-dong^{1,2}, GAO Fei² and ZHANG Li-wei¹

1 Information Center, Ministry of Science and Technology of the People's Republic of China, Beijing 100862, China

2 School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

Abstract Nowadays, the amount of stored and exchanged information in human society is growing geometrically, and the throughput and real-time performance of data transmission need to be improved. While existing studies of network coding focus on improving throughput, the significant impacts of the real-time performance on multipath transmission in big data networks are ignored. This paper addresses the fastest arrival problem for linear network coding, and proposes a critical path computation algorithm with optimized matrix multiplication to improve the real-time performance. In particular, this paper uses the abstract algebra to analyze the critical path algorithm, constructs the commutative ring for the critical path, and proves the optimal substructure property. Simulation results show that the optimized algorithm significantly reduces the time complexity of critical path computation to $O(n^{2.81} \lg n)$, shortens the propagation delays and improves the real-time performance. The time-cost growth rate based on the Strassen critical path algorithm is significantly lower than the repeated square path algorithm while $n > 6$. Specially, while $n = 12$, the computational complexity based on the Strassen critical path algorithm is approximately $2/3$ compared to the repetitive squared critical path algorithm, and the time overhead required is about $1/2$ compared to the latter.

Keywords Critical path, Algebraic structure, Exchange ring, Matrix multiplication, Linear network coding

1 引言

随着无线宽带互联技术日新月异的发展,5G 技术^[1-2]、深空通信^[3-4]和以命名数据网络(Named Data Networking, NDN)^[5-6]为代表的信息中心网络等实际应用都需要高质量的海量数据传输,再加上智能终端的普及,网络用户对高清视频的即时传播需求愈发强烈^[7-8]。但是由于网络节点的计算

能力、传输带宽与存储资源有限,在无线频谱资源日益紧张与贫乏的今天,如何满足海量数据的高效传输是一个亟待解决的问题。网络时延由传输时延和传播时延两部分组成,它是衡量海量数据即时高效传输的一个重要指标。网络编码能够逼近网络吞吐量的极限,并提升数据传输的有效性^[9-10],可被用在信息中心网络等应用中来改善数据的实时传输。但是在基于网络编码操作的信息中心网络中,目前的网络编码技术

到稿日期:2019-08-06 返修日期:2019-11-04 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金项目(61271258)

This work was supported by the National Natural Science Foundation of China (61271258).

通信作者:高飞(gaofei@bit.edu.cn)

一般只针对传输时延进行优化设计^[11-14],缺少对传播时延的考虑。由于传播时延在空间网络中影响巨大,在以大文件通信为主的信息中心网络中无法忽略,因此为了保证大数据传输的实时性和有效性,需要基于网络编码来研究针对传播时延优化的高效传输算法。

本文研究有向无环图(Directed Acyclic Graph, DAG)中线性网络编码操作下的传播时延与关键路径等效模型,并提出相关优化算法以计算传输时延。网络可以看作 DAG 图,而数据传输的网络流路由则是一种基于网络拓扑结构的选路算法。因为各条路径上的数据处理都完成才标志整个传输任务结束,所以整个传输过程所需的时间取决于源点与汇点间具有最长传播时延的路径即关键路径的长度,而关键路径的长度在线性网络编码操作下即表现为传播时延。因此,研究 DAG 图中线性网络编码操作下的关键路径计算问题并进行优化,对提升网络容量、提高数据传输的实时性及降低传播时延有着非常重要的意义。

2 线性网络编码下关键路径的代数结构

DAG 图中基于网络拓扑的选路算法,比如最短路径算法^[15-17],是在一个闭合的线性空间里进行计算,因此首先要研究此关键路径所具有的代数结构。图中的节点在某种意义上可以看作编码器,可以对接收到的信息数据进行处理;图中的边可以看作信息数据传输的路径,而边上相应的权重则代表数据在此边上传递时所需要的时间即传播时延。网络中节点的入度大于或等于 2 时,就默认此节点对接收到的所有信息进行线性网络编码操作,如图 1 所示。因此,无权图中的最长简单路径问题变成了网络编码场景下对最长传输路径求最优解的问题,而找出最长时延路径可看成关键路径的计算问题。

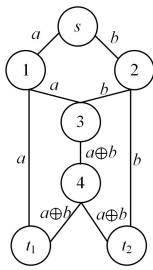


图 1 蝶形网络中的网络编码

Fig. 1 Network coding in butterfly network

定义 1^[18] 设 $G(V, E, \omega)$ 为 n 阶带权 DAG 图,其中入度为 0 的节点称为源点,而边出度为 0 的节点称为汇点,边 $e_{ij} = \langle v_i, v_j \rangle$ 上所赋的权值记为 ω_{ij} ,它的意义为活动的持续时间或者完成此阶段任务所需的代价。

定义 2^[18] 设 n 阶带权 DAG 图 $G(V, E, \omega)$ 包含 n 个节点,其按偏序排列为 $\langle v_1, v_2, \dots, v_n \rangle$,则源点为 v_1 ,汇点为 v_n 。

定义 3^[18] 在带权 DAG 图 $G(V, E, \omega)$ 中,若从节点 v_i 出发,沿着一些边经过某些节点 $v_{p1}, v_{p2}, \dots, v_{pm}$ 到达节点 v_j ,则称节点序列 $\langle v_i, v_{p1}, v_{p2}, \dots, v_{pm}, v_j \rangle$ 为从节点 v_i 到 v_j 的一条路径,其中 $\langle v_i, v_{p1} \rangle, \langle v_{p1}, v_{p2} \rangle, \dots, \langle v_{pm}, v_j \rangle$ 图 G 中的边。

定义 4^[18] 若一条路径上的节点序列号相互之间无交集,则此条路径为基本路径。

在图 2 所示的线性网络结构中,假设信道的传播时延为一个单位时间,则最后一个接收节点的时延 $delay = n - 1$,其中 n 为节点个数;而图 2 所示的 Y-型结构中,信息 a 和 b 经过网络路由同时到达节点 A 和 B 。考虑节点 T 的传播时延,若边 AT 和 BT 相应的权重分别为 1 个和 2 个时间单位,则最后节点 T 完整接收所有信息所需的时延为 2 个时间单位,取决于最大权重的边,因为接收节点需要等待所有的信息完全到达之后才能对信息进行处理。因此,在需要进行线性网络编码操作的 Y-型网络结构中,传播时延问题可以通过关键路径建模为式(1)^[19]。传输时延是指发射机把数据包发向传输介质所需要的时间,取决于发射机的发射速率。传播时延则是指电磁信号在传输介质中传播的时间,取决于电磁波在介质中的传输速率。带宽越大,传输时延越小;距离越长,传播时延越大。

$$\begin{cases} d(T) = \max\{d(AT), d(BT)\} + 1 \\ d(v) = 0; v \in s \end{cases} \quad (1)$$

此问题可进一步写成式(2)^[19]:

$$d(u) = \max\{d(v)\} + 1; v \in \sigma(u) \quad (2)$$

其中, $\sigma(u)$ 表示进行线性网络编码操作的节点 u 的所有上一级节点的集合。这是一类动态规划问题,具有动态规划解,即将无权最长简单路径问题变成了基于线性网络编码的最长路径求最优解的问题。定理 1 证明了此关键路径问题具有最优子结构,定理 2 证明了此代数结构可以看成由集合 $d(u)$ 和二元运算 \cdot 构成的系统 R ,其满足交换环的性质。由此证明了此关键路径与传播时延的等效模型的正确性,即可以采用关键路径所具有的代数结构来计算 DAG 图中线性网络编码下的传播时延问题。

定理 1^[19] $d(u) = \max\{d(v)\} + 1, v \in \sigma(u)$,具有最优子结构。

证明略。

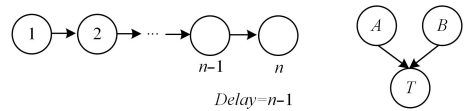


图 2 线性网络与 Y-型网络结构的时延对比

Fig. 2 Delay comparison with linear and Y-shape network

定理 2 线性网络编码操作下的关键路径结构 R 是一个满足交换环的代数结构。

证明:(1)对于所有的 $a, b \in v$,有 $d(T_1) \oplus d(T_2) \in R$ 以及 $d(T_1) \otimes d(T_2) \in R$ 。

(2)分别取节点 $a, b \in v$,有:

$$\begin{aligned} & d(T_1) \oplus d(T_2) \\ &= (\max\{d(a)\} + 1) \oplus (\max\{d(b)\} + 1) \\ &= (\max\{d(b)\} + 1) \oplus (\max\{d(a)\} + 1) \\ &= d(T_2) \oplus d(T_1) \end{aligned}$$

(3)对于 $a, b, c \in v$,有:

$$d(T_1) \oplus (d(T_2) \oplus d(T_3))$$

$$\begin{aligned}
&= (\max\{d(a)\} + 1) \oplus \{(\max\{d(b)\} + 1) \oplus (\max\{d(c)\} + 1)\} \\
&= \{(\max\{d(a)\} + 1) \oplus (\max\{d(b)\} + 1)\} \oplus (\max\{d(c)\} + 1) \\
&= d(T_1) \oplus d(T_2) \oplus d(T_3)
\end{aligned}$$

(4) 当 $v \in s$ 时 $d(v) = d(s) = 0$, 则 $d(s)$ 为系统中的零元素, 即加法单位元。对于任意的 $a \in v$, 有:

$$\begin{aligned}
d(T_1) &= \max\{d(a) \oplus d(s)\} + 1 \\
&= \max\{d(a) \oplus 0\} + 1 \\
&= \max\{d(a)\} + 1 \\
&= 0 \oplus d(T_1) \\
&= d(s) \oplus d(T_1)
\end{aligned}$$

(5) 对于 $a, b \in R$, 有:

$$\begin{aligned}
d(T_1) \otimes d(T_2) &= (\max\{d(a)\} + 1) \otimes (\max\{d(b)\} + 1) \\
&= \max\{d(a)\} \otimes (\max\{d(b)\} + 1) \oplus (\max\{d(b)\} + 1) \\
&= (\max\{d(b)\} + 1) \otimes (\max\{d(a)\} + 1) \\
&= d(T_2) \otimes d(T_1)
\end{aligned}$$

(6) 对于 $a, b, c \in R$, 有:

$$\begin{aligned}
d(T_1) \otimes (d(T_2) \otimes d(T_3)) &= (\max\{d(a)\} + 1) \otimes \{(\max\{d(b)\} + 1) \otimes (\max\{d(c)\} + 1)\} \\
&= (\max\{d(a)\} \otimes \max\{d(b)\}) \otimes (\max\{d(c)\} + 1) \oplus \max\{d(c)\} \otimes (\max\{d(a)\} \oplus \max\{d(b)\}) \oplus (\max\{d(a)\} \oplus \max\{d(b)\}) \oplus (\max\{d(c)\} + 1) \\
&= (\max\{d(c)\} + 1) \otimes \{(\max\{d(a)\} \otimes \max\{d(b)\}) \oplus \max\{d(a)\} \oplus (\max\{d(b)\} + 1)\} \\
&= (d(T_1) \otimes d(T_2)) \otimes d(T_3)
\end{aligned}$$

(7) 当 $v \in s$ 时, 设信源 s 为第 0 级节点, 那么第 1 级节点的时延可以表示为 $d(v_1) = \max\{d(s)\} + 1 = 1$, 则 $d(v_1)$ 可以看作系统的乘法单位元 e , 且对于任意的 $a \in R$, $d(v_1) \otimes d(a) = d(a)$ 成立。

(8) 对于 $a, b, c \in R$, 有:

$$\begin{aligned}
d(T_1) \otimes (d(T_2) \oplus d(T_3)) &= (\max\{d(a)\} + 1) \otimes \{(\max\{d(b)\} + 1) \oplus (\max\{d(c)\} + 1)\} \\
&= (\max\{d(a)\} \otimes \max\{d(b)\}) \oplus \max\{d(a)\} \oplus \max\{d(b)\} + 1 \oplus (\max\{d(a)\} \otimes \max\{d(c)\}) \oplus \max\{d(a)\} \oplus \max\{d(c)\} + 1 \\
&= (\max\{d(a)\} + 1) \otimes (\max\{d(b)\} + 1) \oplus (\max\{d(a)\} + 1) \otimes (\max\{d(c)\} + 1) \\
&= (d(T_1) \otimes d(T_2)) \oplus (d(T_1) \otimes d(T_3))
\end{aligned}$$

同理可证:

$$\begin{aligned}
(d(T_1) \oplus d(T_2)) \otimes d(T_3) &= (d(T_1) \otimes d(T_3)) \oplus (d(T_2) \otimes d(T_3))
\end{aligned}$$

综上所述, $R(\oplus, \otimes, 0, 1)$ 是一个交换环。定理 2 得证。

最短路径的最优子结构的性质^[18]: 如果 $P(i, j) = \{v_i, \dots, v_k, \dots, v_s, \dots, v_j\}$ 是从节点 i 到 j 的最短路径, k 和 s 分

别是这条路径上的中间顶点, 那么 $P(k, s)$ 必定是从节点 k 到 s 的最短路径。

联合定理 1、定理 2 以及最优子结构性可知, 基于线性网络编码的最长路径求最优解的问题与单源最短路径问题具有相同的最优子结构, 即 DAG 图中线性网络编码操作下的关键路径代数结构等价于单源最短路径的代数结构, 因此可以利用此代数结构 $R(\oplus, \otimes, 0, 1)$ 求解线性网络编码规则下关键路径的计算问题。

3 线性网络编码下的关键路径算法

本节基于关键路径具有交换环的代数结构提出了关键路径计算算法, 并基于相关思想对其进行了优化。令 $G(V, E, w)$ 是一个节点个数为 $|V| = n$ 且赋权边的权值取自代数结构 $R(\oplus, \otimes, 0, 1)$ 的 DAG 图, 其中 $P \rightarrow Q$ 表示路径 P 经过节点 v 与路径 Q 相关联。对于节点 x_1, x_2, \dots, x_k , 若任意节点之间存在一条边 (x_i, x_{i+1}) ($i = 1, \dots, k-1$), 则路径可以表示为 $x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_k$, 其连续地经过每一个节点 x_i 。如果路径长度为 0, 则其权重为 0, 当路径 P 仅包含一条边 (x, y) 时, 相应的权重表示为 $w(x, y)$ ^[20]。通过递推可以得到当路径长度 $t \geq 2$ 时, 权重表示为 $w(P) = (w(x_0 \rightarrow \dots \rightarrow x_{t-1})) \otimes w(x_{t-1}, x_t)$ 。令 P_{xy} 表示节点 x 和 y 之间所有有限长路径的集合, 因为最长路径求解问题即关键路径的长度在线性网络编码操作下可以等效于计算传播时延, 所以 $w(P)$ 代表路径的时延。

定义 5 节点 x 和 y 之间关键路径的长度为:

$$D[x, y] = \oplus_{P \in P_{xy}} w(P)$$

定义 6 $n \times n$ 邻接矩阵为:

$$A[i, j] := \begin{cases} w(i, j), & i \neq j \\ 1, & \text{otherwise} \end{cases}$$

其中, $w \in R; i, j \in n$ 。

定义 7 $R(\oplus, \otimes, 0, 1)$ 上的 $n \times n$ 单位阵为:

$$I[i, j] := \begin{cases} 0, & i \neq j \\ 1, & \text{otherwise} \end{cases}$$

令 $M(n, R)$ 是 R 上 $n \times n$ 矩阵的集合, 那么对于其中任意两个矩阵 $A, B \in M(n, R)$, 有 $(A \otimes B)[i, j] = \oplus_{k \in [n]} (A[i, k] \otimes B[k, j])$ ^[20]。令 P'_{xy} 为图 $G(V, E, w)$ 中从节点 x 到 y 的路径长度最长为 t 的所有路径集合, 其中 $t \geq 2$ 。

定义 8 路径的权重矩阵为:

$$D^t[x, y] = \oplus_{P \in P'_{xy}} w(P)$$

定义 9 (所有节点对关键路径问题的递归解) 设 D^t_{xy} 为从节点 x 到 y 的至多包含 t 条边的任意路径中的最大权重关键路径矩阵, 其中 $D^{(1)} = (w_{ij}) = A$ 。当 $t \geq 1$ 时, D^t_{xy} 是 $D^{(t-1)}$ 的最大值和从 x 到 y 最多由 t 条边组成的任意路径的最大权重矩阵进行递归定义而得到的, 其中 $D^{(t-1)}$ 是从 x 到 y 最多由 $t-1$ 条边所组成的最长路径的权重矩阵, 具体如式 (3) 所示。

$$\begin{aligned}
D^t_{xy} &= \max(D^{(t-1)}_{xy}, \max_{1 \leq k \leq n} \{D^{(t-1)}_{ik} \oplus w_{ky}\}) \\
&= \max_{1 \leq k \leq n} \{D^{(t-1)}_{ik} \oplus w_{ky}\} \quad (3)
\end{aligned}$$

$$\text{其中, } \mathbf{D}_{xy}^{(i)} = \begin{cases} 0, & i=j \\ -\infty, & i \neq j \end{cases}$$

定理 3 证明了关键路径具有交换环的代数结构,此结构的性质可以在一个封闭的线性空间中通过计算验证得到。本文基于此提出了关键路径计算算法。

定理 3 给定带权重的有向无环图 $G(V, E, \omega)$ 和权重函数 $\omega: E \rightarrow R$, 令 $R(\oplus, \otimes, 0, 1)$ 是一个满足路径条件的代数路径结构, 那么对任意定义在 R 上的 DAG 图中进行线性网络编码操作的关键路径问题, 等价于此代数结构 R 上利用相应的权矩阵进行代数运算求解, 其中 $\otimes \rightarrow +, \oplus \rightarrow \max$, 即关键路径问题可以在交换环 $R' = (R, \max, +)$ 上通过“相加取大”的权矩阵乘积运算进行计算。若不同级的节点之间没有直接相连的边, 则相应的权值为负无穷, 而且带权 DAG 图 $G(V, E, \omega)$ 的关键路径矩阵 \mathbf{D} 能够通过此代数结构 R 被定义。

证明: 假定计算矩阵乘积 $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, 其中 \mathbf{A} 和 \mathbf{B} 均为 $n \times n$ 的矩阵。若在式 (3) 中进行 $\mathbf{D}^{(i-1)} \rightarrow a; \omega \rightarrow b; \mathbf{D}^{(i)} \rightarrow c; \max \rightarrow +; + \rightarrow \cdot$ 的替换, 将得到式 (4), 那么对于 $i, j = 1, 2, \dots, n$, 可通过式 (4) 计算 c_{ij} , 即证明了该算法与矩阵乘法的关系。

$$c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj} \quad (4)$$

定理 3 得证。

关键路径算法的描述如下。

第 1 步 确定给定网络拓扑抽象出的 DAG 图中节点的层次序, 其中信源为第 1 级, 信宿为第 n 级;

第 2 步 按照图中节点的偏序关系求出节点之间的权矩阵;

第 3 步 根据定理 3 求出网络编码系统中关键路径的时延;

第 4 步 依据以上步骤写出此关键路径上的节点。

该算法的核心伪代码如算法 1 所示, 该伪代码可以在给定 \mathbf{A} 和 $\mathbf{D}^{(i-1)}$ 的前提下计算出 $\mathbf{D}^{(i)}$ 。

算法 1 Critical Path Algorithm (\mathbf{D}, \mathbf{A})

1. $n = \mathbf{D}$. rows
2. let $\mathbf{D}' = (\omega_{ij}')$ be a new $n \times n$ matrix
3. for $i = 1$ to n
4. for $j = 1$ to n
5. $\omega_{ij}' = -\infty$
6. for $k = 1$ to n
7. $\mathbf{D}_{ij}' = \max(\mathbf{D}_{ij}', \mathbf{D}_{ik} + \omega_{kj})$
8. return \mathbf{D}'

由于有 3 层嵌套的 for 循环操作, 因此该核心算法的时间复杂度为 $O(n^3)$, 另根据以上分析可知 $\mathbf{D}^{(i)}$ 外还有一层嵌套, 因此算法 1 总的时间复杂度为 $O(n^4)$ 。

4 关键路径算法理论验证模型的分析

本节说明在基于线性网络编码的带权 DAG 图中求解关键路径时延问题的应用, 由此证明传播时延与关键路径的等效模型建模正确, 且此关键路径方法有效。

例 1 给定图 3 所示的带权 DAG 图的网络结构, 图中某

些中间节点的入度大于或等于 2, 满足线性网络编码操作, 边上的权重表示时延。假设从信源 S 发送信息 a 和 b 到信宿 T , 需要求出最后完整接收所有信息的时延。

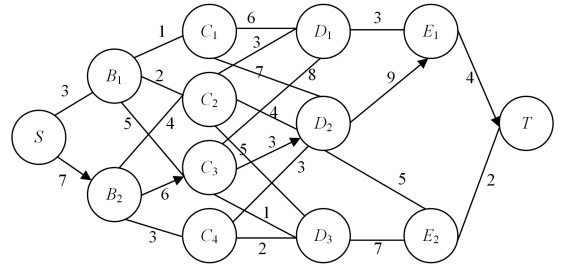


图 3 DAG 网络中的关键路径
Fig. 3 Critical path in DAG network

求解信宿 T 接收全部信息所需要的时延, 即找出带权 DAG 图中的关键路径, 采用定理 3 给出的算法进行求解。

$S \rightarrow B$ 的权矩阵为: $[3 \ 7]$; $B \rightarrow C$ 的权矩阵为:

$$\begin{bmatrix} 1 & 2 & 5 & -\infty \\ -\infty & 4 & 6 & 3 \end{bmatrix}, \text{ 那么 } S \rightarrow C \text{ 的权矩阵为:}$$

$$[3 \ 7] \begin{bmatrix} 1 & 2 & 5 & -\infty \\ -\infty & 4 & 6 & 3 \end{bmatrix} = [4 \ 11 \ 13 \ 10]$$

$S \rightarrow D$ 的权矩阵为:

$$[4 \ 11 \ 13 \ 10] \begin{bmatrix} 6 & 7 & -\infty \\ 3 & 4 & 5 \\ 8 & 3 & 1 \\ -\infty & 3 & 2 \end{bmatrix} = [21 \ 16 \ 16]$$

$S \rightarrow E$ 的权矩阵为:

$$[21 \ 16 \ 16] \begin{bmatrix} 3 & -\infty \\ 9 & 5 \\ -\infty & 7 \end{bmatrix} = [25 \ 23]$$

$S \rightarrow T$ 的权矩阵为:

$$[25 \ 23] \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 29$$

因此, 该网络关键路径的时延为 29 个时间单位, 相应的关键路径为 $S \rightarrow B_2 \rightarrow C_3 \rightarrow D_2 \rightarrow E_1 \rightarrow T$ 。

5 关键路径算法的优化

当网络拓扑结构较复杂时, 拓扑矩阵的维度会较大, 对此通常把大维度的矩阵进行分割计算, 采用分治法渐进地将各个分块小矩阵相乘, 以此对算法进行优化。因此, 将矩阵分块是实现大规模矩阵乘法的有效途径^[21]。关键路径算法是基于朴素矩阵乘法的思想, 总的算法时间效率为 $O(n^4)$, 采用重复平方技术能够将时间复杂度降至 $O(n^3 \lg n)$ 。由于 Strassen 递归算法是分治法的一种典型应用, 因此最终可以通过此思想将算法的运行时间降低到 $O(n^{\lg 7} \lg n) = O(n^{2.81} \lg n)$ 。

首先, 目标是计算矩阵 $\mathbf{D}^{(n-1)}$, 并不是要计算出所有的 $\mathbf{D}^{(m)}$, 因传统朴素矩阵乘法是相关的, 故可以仅用 $\lceil \lg(n-1) \rceil$ 个矩阵乘积来计算 $\mathbf{D}^{(n-1)}$ 。具体的实现过程为: $\mathbf{D}^{(1)} = \mathbf{W}$, $\mathbf{D}^{(2)} = \mathbf{W}^2 = \mathbf{W} \cdot \mathbf{W}$, $\mathbf{D}^{(4)} = \mathbf{W}^4 = \mathbf{W}^2 \cdot \mathbf{W}^2$, \dots , $\mathbf{D}^{(2^{\lceil \lg(n-1) \rceil})} = \mathbf{W}^{2^{\lceil \lg(n-1) \rceil}} = \mathbf{W}^{2^{\lceil \lg(n-1) \rceil - 1}} \cdot \mathbf{W}^{2^{\lceil \lg(n-1) \rceil - 1}}$ 。由于 $2^{\lceil \lg(n-1) \rceil} \geq$

$n-1$, 因此有 $\mathbf{D}^{\lceil 2\lg(n-1) \rceil} = \mathbf{D}^{(n-1)}$ 。接下来根据重复平方的思想给出快速关键路径算法的伪代码(见算法 2), 来计算上述矩阵序列。

算法 2 Faster All Pairs-Critical Path Algorithm (\mathbf{D}, \mathbf{W})

1. $n = \mathbf{W}$. rows
2. $\mathbf{D}^{(1)} = \mathbf{W}$
3. $m = 1$
4. while $m < n - 1$
5. let $\mathbf{D}^{(2^m)}$ be a new $n \times n$ matrix
6. $\mathbf{D}^{(2^m)} = \text{Critical Path Algorithm}(\mathbf{D}^{(m)}, \mathbf{D}^{(m)})$
7. $m = 2m$
8. return $\mathbf{D}^{(m)}$

因为 $\lceil \lg(n-1) \rceil$ 个矩阵中每个矩阵的计算时间为 $O(n^3)$, 所以快速关键路径算法的运行时间为 $O(n^3 \lg n)$ 。

分析定理 3 的证明过程可得到, 求解关键路径即计算矩阵乘积 $\mathbf{C} = \mathbf{A} \times \mathbf{B}$, 其中 \mathbf{A}, \mathbf{B} 和 \mathbf{C} 均为 $n \times n$ 的方阵。首先考虑 $n = 2^k$, 令 $T(n)$ 为两个 $n \times n$ 矩阵相乘所需的算术操作的次数, 它与时间强线性相关。通过引理 1 可得 $T(n)$ 的递归表达式(见式(5))。

引理 1^[21] 两个 $n \times n$ 矩阵的乘积能够通过 7 次递归的 $(n/2) \times (n/2)$ 的矩阵乘法运算和 $\Theta(n^2)$ 次标量加法或减法运算得到, 其中矩阵的元素均从一个任意的环中进行选择且 $n \geq 2$ 。

$$T(n) = 7T(n/2) + \Theta(n^2) = \Theta(n^{\lg 7}) = O(n^{2.81}) \quad (5)$$

证明略。

下面基于 Strassen 的思想给出关键路径算法 (Strassen All Pairs-Critical Path Algorithm) 的步骤。

1) 将 \mathbf{A}, \mathbf{B} 和 \mathbf{C} 都划分为 4 个 $(n/2) \times (n/2)$ 的子矩阵, 等式 $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ 可以改写为式(6):

$$\begin{bmatrix} r & s \\ t & u \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \quad (6)$$

2) 运用 $\Theta(n^2)$ 次标量加法和减法运算, 计算出 14 个 $(n/2) \times (n/2)$ 的子矩阵 $\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{B}_2, \dots, \mathbf{A}_7, \mathbf{B}_7$ 。

3) 递归计算出 7 个矩阵的乘积 $\mathbf{P}_i = \mathbf{A}_i \mathbf{B}_i, i = 1, 2, \dots, 7$ 。

4) 仅使用 $\Theta(n^2)$ 次标量加法与减法运算, 对 \mathbf{P}_i 矩阵的各种组合进行求和或求差运算, 从而得到目标矩阵 \mathbf{C} 的 4 个子矩阵 r, s, t, u 。

上述过程满足递归式(5)。同样地, 因为 $\lceil \lg(n-1) \rceil$ 个矩阵中每个矩阵的计算时间为 $O(n^{2.81})$, 所以最后可以将算法的总运行时间降低到 $O(n^{2.81} \lg n)$ 。随着 n 的变大, 基于 Strassen 思想的关键路径算法在网络中应用的性能将大大提升。

6 仿真结果分析

表 1 对比了采用不同思想优化的关键路径算法的计算复杂度, 其中 n 代表网络节点个数。图 4 给出了在不同网络拓扑结构下采用不同思想的算法的传播时延。通过比较计算量可知, 使用基于重复平方思想的关键路径算法虽然降低了复杂度, 但是优势并不明显。而基于 Strassen 思想优化的关键路径算法在 n 较大时减少的计算量比较突出; 特别地, 当

$n=16$ 时, 基于 Strassen 思想的关键路径算法的计算量仅相当于朴素矩阵乘法的 1/27。

表 1 不同思想的算法下矩阵相乘的计算量比较

Table 1 Comparison of calculations of matrix multiplication under different algorithms

算法	$n=2$	$n=4$	$n=6$	$n=8$	$n=10$	$n=12$	$n=14$	$n=16$
朴素矩阵乘法 $O(n^4)$	16	256	1 296	4 096	10 000	20 736	38 416	65 536
重复平方法 $O(n^3)$	8	64	216	512	1 000	1 728	2 744	4 096
基于 Strassen $O(n^{2.52})$	7	49	154	345	646	1 078	1 662	2 419

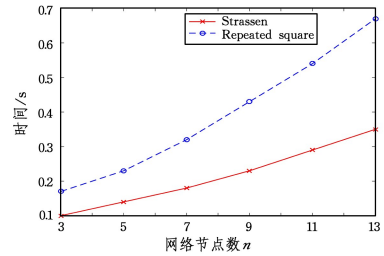


图 4 节点个数与不同思想的算法的传播时延

Fig. 4 Delay of different algorithms with network nodes

图 4 给出了节点个数 n 与数据包传播时延的关系。首先, 时延与矩阵大小成正比, 随着 n 的不断增大, 时延也相应增加。从图中可以看出, 在 $n \leq 5$ 时, 基于 Strassen 与基于重复平方的算法的运算性能已经有了一定的差异。但当矩阵规模达到 $n > 6$ 之后, 基于 Strassen 思想的关键路径算法的时间开销增长速率显著低于基于重复平方思想的算法的时间开销增长速率; 特别地, 当 $n = 12$ 时, 基于 Strassen 的关键路径算法的计算量约是基于重复平方的关键路径算法计算量的 2/3, 而其所需的时间开销约为后者的 1/2。因此, 对于规模越大的并行矩阵运算, 基于 Strassen 思想的关键路径算法的性能优势越明显, 这对于当前以海量数据分发为主的信息中心网络的传输性能提升具有重要意义。仿真实验验证了所提模型的正确性, 此关键路径算法能够以很低的时间复杂度有效地获取基于线性网络编码的有向无环图中传播时延最长的路径, 提高了数据传输效率和编码增益, 缩短了通信时延, 使得数据能更快地到达接收者。

结束语 本文针对网络时延性能研究了 DAG 图中线性网络编码下的关键路径计算问题, 提出了传播时延与关键路径的等效模型。本文突破了现有研究的局限性, 考虑了传播时延, 从而得到了更科学的时延计算方法。该方法具有清晰的代数结构, 更便于理论分析。本文进而设计了一种基于交换环的矩阵乘法的关键路径算法并对其进行了优化, 另通过应用举例说明了该算法具有直观性强、计算简单的特点。相比于网络中传统的找路思想, 该算法利用抽象代数与图论的相关知识对 DAG 图中的代数路径结构进行了研究及计算, 将关键路径计算问题转化为在一个封闭的代数空间中, 利用矩阵乘积的“相加取大”代数运算操作进行求解的简单有向路径问题, 减少了冗余计算, 提高了算法效率, 进一步优化了

大数据网络中应用网络编码的限制条件,扩展了网络编码技术研究 and 应用的范围,保证了网络数据的高效传输。仿真结果显示,利用代数结构计算和优化线性网络编码的关键路径,可以增强文件传输的实时性,提升网络容量,改善网络的传输性能。

参 考 文 献

- [1] AKPAKWU G A, SILVA B J, HANCKE G P, et al. A Survey on 5G Networks for the Internet of Things: Communication Technologies and Challenges [J]. *IEEE Access*, 2018, 6: 3619-3647.
- [2] FANG D F, QIAN Y, HU Q Y. Security for 5G Mobile Wireless Networks [J]. *IEEE Access*, 2018, 6: 4850-4874.
- [3] CUI J X, DONG G L, LI H T, et al. Prospect of Application of Physical Layer Network Coding in Deep Space Communication [C]//2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS). 2016: 131-134.
- [4] JARZYNA M, ZWOLINSKI W, JACHURA M, et al. Optimizing deep-space optical communication under power constraints [C]//Free-Space Laser Communication and Atmospheric Propagation XXX. International Society for Optics and Photonics, 2018.
- [5] VAN J, SMETTERS D K, THORNTON J D, et al. Networking Named Content [C]//CoNEXT'09 Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies. New York: ACM, 2009: 1-12.
- [6] CAO J X, PEI D, ZHANG X P, et al. Fetching Popular Data from the Nearest Replica in NDN [C]//2016 25th International Conference on Computer Communication and Networks (ICCCN). 2016: 1-9.
- [7] ZHAO Z F, TAN X B, SU J X. Adaptive traffic scheduling via network coding in NDN [C]//2016 35th Chinese Control Conference (CCC). 2016: 7200-7205.
- [8] AUBRY E, SILVERSTON T, CHRISMENT I. Green growth in NDN: Deployment of content stores [C]//2016 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN). 2016: 1-6.
- [9] ALSHWEDE R, CAI N, LI S Y R, et al. Network information flow [J]. *IEEE Transactions on Information Theory*, 2000, 46(4): 1204-1216.
- [10] HO T, MÉDARD M, KOETTER R, et al. A Random Linear Network Coding Approach to Multicast [J]. *IEEE Transactions on Information Theory*, 2006, 52(10): 4413-4430.
- [11] HU G, XU K, XU Y. Deadline-Aware Opportunistic Network Coding for Multi-Relay-Aided Single-Source Single-Destination Network [J]. *IEEE Communications Letters*, 2017, 21(10): 2282-2285.
- [12] ZHU J. Exploiting Opportunistic Network Coding for Improving Wireless Reliability Against Co-Channel Interference [J]. *IEEE Transactions on Industrial Informatics*, 2016, 12(5): 1692-1701.
- [13] ZOHDY M, ELBATT T, NAFIE M. Maximum Throughput Opportunistic Network Coding in Two-Way Relay Networks [C]//IEEE International Conference on Communication. 2015: 913-918.
- [14] AMIRI M M, GUNDUZ D. Fundamental Limits of Coded Caching: Improved Delivery Rate-Cache Capacity Tradeoff [J]. *IEEE Transactions on Communications*, 2017, 65(2): 806-815.
- [15] JOHARI S, OJHA A. Quickest path problems in stochastic-flow networks with time constraint: A fast and reliable solution [C]//2013 3rd IEEE International Advance Computing Conference (IACC). 2013: 322-325.
- [16] SEGARRA S, MARQUES A G, RIBEIRO A. Optimal Graph-Filter Design and Applications to Distributed Linear Network Operators [J]. *IEEE Transactions on Signal Processing*, 2017, 65(15): 4117-4131.
- [17] ETZION T, STORME L. Galois geometries and coding theory [J]. *Designs, Codes and Cryptography*, 2016, 78(1): 311-350.
- [18] CORMEN T H, LEISERSON C E, et al. Introduction to Algorithm [M]. MIT Press, 2005.
- [19] HAN X D, GAO F, WANG A H. Network Coding Transmission Performance Under the NDN Protocol [J]. *Transactions of Beijing Institute of Technology*, 2017, 37(5): 506-510.
- [20] VASSILEVSKA V. Efficient Algorithms for Path Problems [D]. Pittsburgh: Carnegie Mellon University, 2008.
- [21] AHO A V, HOPCROFT J E, ULLMAN J D. The Design and Analysis of Computer Algorithms [M]. Addison-Wesley Publishing Company, 1974.



HAN Xiao-dong, born in 1983, Ph.D., engineer. Her main research interests include network coding and network security.



GAO Fei, born in 1959, Ph.D., professor. Her main research interests include wireless communication technology, image signal processing and network security.