[36] Neely MJ. Stochastic Network Optimization with Application to Communication and Queueing Systems. Morgan and Claypool, 2010.

[37] Neely MJ. Opportunistic scheduling with worst case delay guarantees in single and multi-hop networks. In: Proc. of the INFOCOM. IEEE, 2011. 1728−1736. [doi: 10.1109/INFCOM.2011.5934971]

[38] Arlitt M, Jin T. A workload characterization study of the 1998 world cup Web site. IEEE Network, 2000,14(3):30–37. [doi: 10.1109/65.844498]

[39] Umasstracerepository. 2014. http://traces.cs.umass.edu/index.php/Network/Network

[40] Where Amazon's data centers are located. http://www.datacenterknowledge.com/archives/2008/11/18/where-amazons-data-centers-are-located/

[41] Gpsspg. http://www.gpsspg.com/distance.htm

**附中文参考文献:**

[7] 印鉴,王智圣,李琪,苏伟杰.基于大规模隐式反馈的个性化推荐.软件学报,2014,25(9):1953−1966. http://www.jos.org.cn/1000-9825/4648.html [doi: 10.13328/j.cnki.jos.004648]

[8] 韩乐,黎铭.基于代价敏感多标记学习的开源软件分类.软件学报,2014,25(9):1982−1991. http://www.jos.org.cn/1000-9825/4639.html [doi: 10.13328/j.cnki.jos.004639]

[9] 谢娟英,高红超.基于统计相关性与 K-means 的区分基因子集选择算法.软件学报,2014,25(9):2050−2075. http://www.jos.org.cn/1000-9825/4644.html [doi: 10.13328/j.cnki.jos.004644]

[11] 徐计,王国胤,于洪.基于粒计算的大数据处理.计算机学报,2015,38(8):1497−1517.

[12] 丁有伟,秦小麟,刘亮,王涛春.一种异构集群中能量高效的大数据处理算法.计算机研究与发展,2015,52(2):377−390.

[13] 李玮,张大方,黄昆,谢鲲.面向大数据处理的高精度多维计数布鲁姆过滤器.电子学报,2015,43(4):752−657.

[14] 卢志茂,冯进玫,范冬梅,杨朋,田野.面向大数据处理的划分聚类新方法.系统工程与电子技术,2014,36(5):1010−1015.

## 附录 A:漂移惩罚上界推导

证明:根据$(\max[x-y,0]+z)^2 \leqslant x^2+y^2+z^2+x(z-y)$,对向量$\Theta(t)=[\boldsymbol{Z}(t),\boldsymbol{H}(t)]$则有:

$$\begin{aligned} Z_d(t+1)^2 - Z_d(t)^2 &\leqslant \left\{1_{(H(t)>0)}\left(\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right) - 1_{(H(t)=0)}\sum_{k\in\mathcal{K}} n_d^{k,\max}v_k\right\} + \\ &\quad 2Z_d(t)\left\{1_{(H(t)>0)}\left(\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right) - 1_{(H(t)=0)}\sum_{k\in\mathcal{K}} n_d^{k,\max}v_k\right\} \\ &\leqslant (\varepsilon_d)^2 + \left(\sum_{k\in\mathcal{K}} n_d^{k,\max}v_k\right)^2 + 2Z_d(t)\left(\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right) \end{aligned} \tag{A.1}$$

$$H_d(t+1)^2 - H_d(t)^2 \leqslant \left(\sum_{k\in\mathcal{K}} n_d^k(t)v_k\right)^2 + \left(\sum_{r\in\mathcal{R}} \lambda_r^d(t)\right)^2 + 2H_d(t)\left(\sum_{r\in\mathcal{R}} \lambda_r^d(t) - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right) \tag{A.2}$$

由于 $\lambda_r^d(t)$ 和 $n_d^k(t)$ 的最大值分别为 $A_{\max}^r$ 和 $N_d^{k,\max}$,设 $B=\frac{1}{2}\sum_{d\in\mathcal{D}}\left\{2\left(\sum_{k\in\mathcal{K}} N_d^{k,\max}v_k\right)^2 + (\varepsilon_d)^2 + \left(\sum_{r\in\mathcal{R}} A_{\max}^r\right)\right\}$,则可得单时隙李雅普诺夫漂移如下:

$$\begin{aligned} \Delta(\Theta(t)) &= \frac{1}{2}\sum_{d\in\mathcal{D}} \mathbb{E}\{H_d(t+1)^2 - H_d(t)^2 \mid \Theta(t)\} + \frac{1}{2}\sum_{d\in\mathcal{D}} \mathbb{E}\{Z_d(t+1)^2 - Z_d(t)^2 \mid \Theta(t)\} \\ &\leqslant B + \sum_{d\in\mathcal{D}} \mathbb{E}\left\{H_d(t)\left(\sum_{r\in\mathcal{R}} \lambda_r^d(t) - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right)\middle|\Theta(t)\right\} + \sum_{d\in\mathcal{D}} \mathbb{E}\left\{Z_d(t)\left(\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(t)v_k\right)\middle|\Theta(t)\right\} \end{aligned} \tag{A.3}$$

通过在上述表达中加上 $V\cdot\mathbb{E}\{C(t)\}$,我们得到了公式(20)中的漂移惩罚界限. □

## 附录 B:定理 1 的证明

为证明定理 1,我们首先给出如下引理.

**引理 B.1**(**最优随机稳定策略的存在性**). 至少存在一个策略$\pi$,使其选择的可行解 $n_d^{k,\pi}(t),\lambda_d^{k,\pi}(t)$ ($\forall d\in\mathcal{D}$, $r\in\mathcal{R},k\in\mathcal{K},t\in[1,T]$)满足:

$$\begin{aligned}&\mathbb{E}\{C(t)\}=C^*\\&\mathbb{E}\left\{\sum_{r\in\mathcal{R}}\lambda_r^{d,\pi}(t)\right\}\leqslant\mathbb{E}\left\{\sum_{k\in\mathcal{K}}n_d^{k,\pi}(t)v_k\right\}\\&\varepsilon_d\leqslant\mathbb{E}\left\{\sum_{k\in\mathcal{K}}n_d^{k,\pi}(t)v_k\right\}\end{aligned}\tag{A.4}$$

其中,$C^*$是理论成本下界.该引理可利用卡拉特欧多定理证明,详细见文献[36].

根据引理 B.1,我们可如下证明公式(27).

证明:从引理 B.1 可知,必然存在一个常数$\delta$>0,满足:

$$\mathbb{E}\left\{\sum_{r\in\mathcal{R}}\lambda_r^{d,\pi}(t)\right\}\leqslant\mathbb{E}\left\{\sum_{k\in\mathcal{K}}n_d^{k,\pi}(t)v_k\right\}-\delta\tag{A.5}$$

以及

$$\varepsilon_d\leqslant\mathbb{E}\left\{\sum_{k\in\mathcal{K}}n_d^{k,\pi}(t)v_k\right\}-\delta\tag{A.6}$$

由本文算法力求最小化不平等公式(20)的右半部分,通过在各时隙所有可行的决策中选择最优决策,应用引理 B.1,将公式(A.5)和公式(A.6)代入公式(20),我们可得:

$$\Delta(\Theta(t))+V\cdot\mathbb{E}\{C(t)\mid\Theta(t)\}\leqslant B+VC^*-\delta\sum_{d\in\mathcal{D}}\mathbb{E}\{H_d(t)\}-\delta\sum_{d\in\mathcal{D}}\mathbb{E}\{Z_d(t)\}\tag{A.7}$$

取公式(A.7)的期望,并应用$\Delta(\Theta(t))=\mathbb{E}\{L(\Theta(t+1))-L(\Theta(t))|\Theta(t)\}$,可得:

$$\mathbb{E}\{L(\Theta(t+1))-L(\Theta(t))\mid\Theta(t)\}+V\cdot\mathbb{E}\{C(t)\mid\Theta(t)\}\leqslant B+VC^*-\delta\sum_{d\in\mathcal{D}}\mathbb{E}\{H_d(t)\}-\delta\sum_{d\in\mathcal{D}}\mathbb{E}\{Z_d(t)\}\tag{A.8}$$

在时隙 $t$=1,…,$T$−1 上应用伸缩和定理(telescoping sum law),并将结果除以 $T$,可得:

$$\frac{\mathbb{E}\{L(\Theta(T))-L(\Theta(0))\}}{T}+\frac{V}{T}\cdot\sum_{t=0}^{T-1}\mathbb{E}\{C(t)\mid\Theta(t)\}\leqslant B+VC^*-\frac{\delta}{T}\sum_{t=0}^{T-1}\sum_{d\in\mathcal{D}}\mathbb{E}\{H_d(t)\}-\frac{\delta}{T}\sum_{t=0}^{T-1}\sum_{d\in\mathcal{D}}\mathbb{E}\{Z_d(t)\}\tag{A.9}$$

对上述公式进行整理,并考虑约束:$L(\Theta(0))=0,H_d(t)\geqslant0,Z_d(t)\geqslant0$,得:

$$\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\{C(t)\}\leqslant C^*+\frac{B}{V}\tag{A.10}$$

对公式(A.10)求极限 $T\to\sim$,可得公式(26). □

## 附录 C:定理 2 的证明

证明:对于 $Z_d(t)$,已知 $Z_d(0)=0<Z_d^{\max}$ .对任意时刻 $t\in[0,T]$:

- 若 $Z_d(t)\leqslant\frac{Vp_d^{\max}}{v_{\min}}$ ,则 $Z_d(t+1)\leqslant\frac{Vp_d^{\max}}{v_{\min}}+\varepsilon_d=Z_d^{\max}$ (因为如公式(14)所示,$Z_d(t)$最大增量为$\varepsilon_d$);
- 若 $Z_d(t)>\frac{Vp_d^{\max}}{v_{\min}}$ ,则 $H_d(t)+Z_d(t)>\frac{Vp_d^{\max}}{v_{\min}}>\frac{Vp_d^k}{v_k}$ ,并且队列将减少 $\sum_{k\in\mathcal{K}}N_d^{k,\max}v_k$ (见公式(26)).

而 $\varepsilon_d<\sum_{k\in\mathcal{K}}N_d^{k,\max}v_k$ ,此时队列 $Z_d(t)$的增量小于减小量,因此 $Z_d(t+1)<Z_d(t)<Z_d^{\max}$ .队列 $Z_d(t)$的界限得证.

类似地,对于队列 $H_d(t)$, $H_d(0)=0<H_d^{\max}$ .对任意时刻 $t$:

- 若 $H_d(t) \leqslant \dfrac{Vp_d^{\max}}{v_{\min}}$,则 $H_d(t+1) \leqslant \dfrac{Vp_d^{\max}}{v_{\min}} + \sum_{r\in\mathcal{R}} A_{\max}^r$ (由公式(13),$H_d(t+1)$最大增量为 $\sum_{r\in\mathcal{R}} A_{\max}^r$ );
- 若 $H_d(t) > \dfrac{Vp_d^{\max}}{v_{\min}}$,则 $H_d(t)+Z_d(t) > \dfrac{Vp_d^{\max}}{v_{\min}} \geqslant \dfrac{Vp_d^k}{v_k}$,由公式(26)可知,$t$+1 时刻队列将减小 $\sum_{k\in\mathcal{K}} N_d^{k,\max} v_k$.

而 $\sum_{r\in\mathcal{R}} A_{\max}^r \leqslant \sum_{k\in\mathcal{K}} N_d^{k,\max} v_k$,即队列的增量小于减小量,因而 $H_d(t+1) < H_d(t) < H_d^{\max}$. □

因此,队列 $H_d(t)$的界限得证.

## 附录 D:定理 3 的证明

证明:若存在$\tau\in[t+1,t+l]$满足 $H_d(\tau)=0$,则在 $t$ 时刻到达的数据将在 $l$ 个时间序列内被处理.若 $H_d(\tau)>0,\tau\in[t+1,t+l]$,根据公式(14),可得:

$$Z_d(\tau+1) = \max\left[Z_d(\tau)+\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(\tau)v_k, 0\right] \geqslant Z_d(\tau)+\varepsilon_d - \sum_{k\in\mathcal{K}} n_d^k(\tau)v_k \tag{A.11}$$

在时间序列$\tau\in[t+1,t+l]$内对以上的不等式求和,可得:

$$Z_d(t+l+1) \geqslant Z_d(t+1) + l\cdot\varepsilon_d - \sum_{\tau=t+1}^{t+l+1}\sum_{k\in\mathcal{K}} n_d^k(\tau)\cdot v_k \tag{A.12}$$

因此,可以推导出 $\sum_{\tau=t+1}^{t+l+1}\sum_{k\in\mathcal{K}} n_d^k(\tau)\cdot v_k \geqslant l\cdot\varepsilon_d - Z_d^{\max}$. 注意:当 $l=[H_d^{\max}+Z_d^{\max}/\varepsilon_d]$ 时,有 $l\cdot\varepsilon_d - Z_d^{\max} = H_d^{\max}$, 则

$$\sum_{\tau=t+1}^{t+l+1}\sum_{k\in\mathcal{K}} n_d^k(\tau)\cdot v_k \geqslant H_d^{\max} \geqslant H_d(t) \tag{A.13}$$

由于队列系统以先进先出方式运行,在 $t$ 时刻到达的数据将比在 $t$ 时刻之后到达的数据先处理.既然 $l$ 个时间序列被处理的总数据量大于 $H_d(t)$,那么所有在 $t$ 时刻到达的数据都将在 $l$ 个时间序列中被处理,即最差的延迟为 $l$ 个时隙. □


**肖文华**(1988-),男,湖南桂阳人,博士生,主要研究领域为云计算,移动云计算中的资源管理问题.

**邵屹杨**(1992-),男,硕士生,主要研究领域为云计算容错管理研究.

**包卫东**(1971-),男,博士,教授,博士生导师,主要研究领域为信息系统.

**陈超**(1990-),男,博士生,主要研究领域为云计算,移动云计算.

**朱晓敏**(1979-),男,博士,副教授,CCF 高级会员,主要研究领域为云计算资源管理,实时系统.

**Jianhong WU** (1964-),男,博士,教授,博士生导师,主要研究领域为微分方程动力学,面向大数据分析的神经网络架构.