Multi-objective Optimization for Computation Offloading in Mobile-edge Computing

报告人: 王琦 导师: 李智勇





论文介绍:

题目:《Multi-objective Optimization for Computation Offloading in Mobile-edge Computing》

会议: 2017 IEEE Symposium on Computers and Communications (ISCC) C类

目标: 能源消耗、执行延迟和价格成本的多目标优化问题, 寻找最优卸载概率和最优传输功率。

贡献:

- (1)以往的文献通常在设计计算卸载方案时会单独考虑**能耗、延迟**或者**利用云资源的成本**,本文首次同时考虑这三个优化的目标。
- (2) 无线信道的传输功率可变(以前为固定值);

边缘的计算能力有限(以前认为无限)



背景

问题:

由于移动设备(MDs)在尺寸、重量、电池寿命、散热等方面的限制, 许多计算密集型和延迟敏感的移动应用程序在智能手机上运行时性能很 差,比如图像处理、象棋游戏等等。

方案:

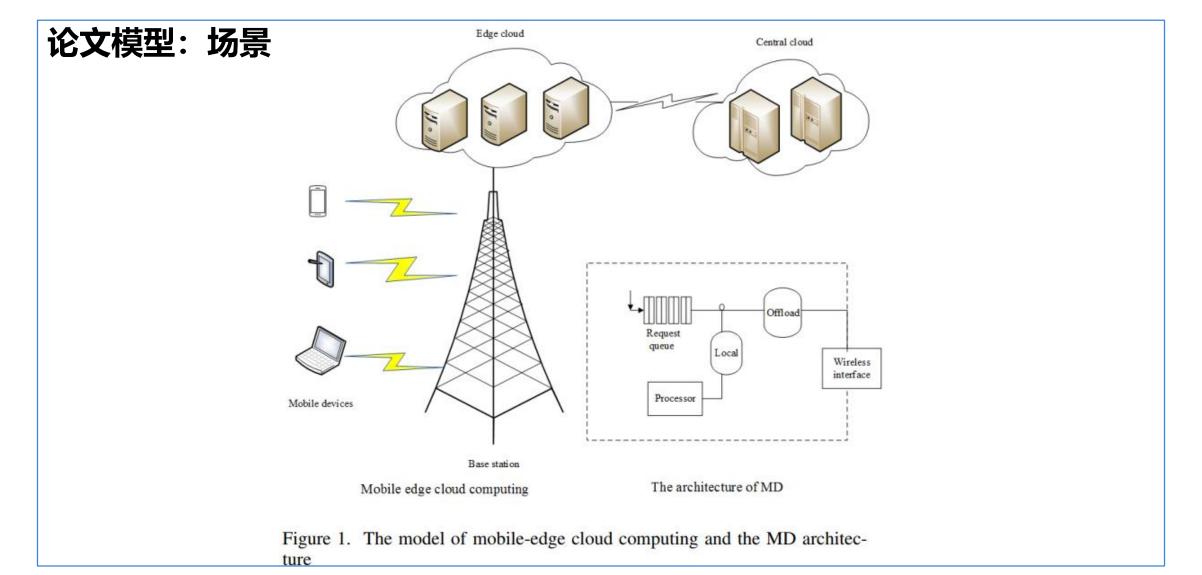
最近的研究表明,移动云计算(MCC)技术通过将计算密集型任务转移到 云上执行,为移动设备克服硬件限制和节约能源提供了一个有效的方法。



前述文献结论:

- (1) 相同大小的请求,传输能量消耗通常小于本地执行能量消耗。
- (2) 相同大小的请求, 卸载执行时间大于本地执行时间。
- (3)移动设备必须在边缘云或中央云中使用的资源付费。







论文模型:设备端

- MD的排队模型为 M/M/1
- 响应时间为: $R = \frac{1/u}{1-\rho}$ where $\rho = \frac{\lambda}{u}$

MDi: 第*i*个移动设备, *i*∈ {1,2,...,N}

 P_i^C : MD i 的卸载概率

 λ_i : 服从泊松分布的平均到达率

 u_i^M : *MD i* 的计算能力

 l_i^M : MDi的CPU使用率

 κ_{i} : 本地计算功率

本地执行的时间:
$$T_i^M(p_i^C) = \frac{1}{u_i^M(1-l_i^M) - (1-p_i^C)\lambda_i}$$
. (1)

本地执行消耗的能量:
$$E_i^M\left(p_i^C\right) = \kappa_i T_i^M\left(p_i^C\right) = \kappa_i \frac{1}{u_i^M\left(1 - l_i^M\right) - \left(1 - p_i^C\right)\lambda_i}$$
, (4)



论文模型: 设备端

W: 信道带宽

 P_i : *MDi* 的传输功率($0 \le P_i \le P_i^{th}$)

上行数据传输率:
$$R_i = W \log_2 \left(1 + \frac{P_i H_{i,s}}{\omega_i + \sum_{j \in N, j \neq i} P_j H_{j,s}} \right)$$
, (2) $H_{j,s}$: 信道增益 ω_i : 背景干扰功率

 θ_i : 每个request的数据大小

上行数据传输时间:
$$T_i^t\left(p_i^C, P_i\right) = \frac{p_i^C \lambda_i \theta_i}{R_i} = \frac{p_i^C \lambda_i \theta_i}{W \log_2\left(1 + \frac{P_i H_{i,s}}{\omega_i + \sum_{j \in N, j \neq i} P_j H_{j,s}}\right)}$$
. (3)

数据传输消耗的能量:

$$E_{i}^{S}\left(p_{i}^{C}, P_{i}\right) = P_{i}T_{i}^{t}\left(p_{i}^{C}\right) = P_{i}\frac{p_{i}^{C}\lambda_{i}\theta_{i}}{R_{i}}$$

$$= \frac{P_{i}p_{i}^{C}\lambda_{i}\theta_{i}}{W\log_{2}\left(1 + \frac{P_{i}H_{i,s}}{\omega_{i} + \sum_{i \in N} \sum_{i \neq i} P_{i}H_{i,s}}\right)}.$$
(5)



论文模型:边缘端

• 边缘端的排队模型为 M/M/c

边缘端占用的服务率:
$$\lambda_{Total}^{M} = \sum_{i=1}^{N} \lambda_{i} p_{i}^{C}$$
. (6)

在边缘端执行的概率:

$$\psi^{C} = \begin{cases} 1, & \lambda_{\text{max}}^{C} \ge \lambda_{Total}^{M}; \\ \frac{\lambda_{\text{max}}^{C}}{\lambda_{Total}^{M}}, & \lambda_{\text{max}}^{C} < \lambda_{Total}^{M}; \end{cases}$$
(7)

实际的服务率:
$$\lambda_p^C = \psi^C \lambda_{Total}^M = \begin{cases} \lambda_{Total}^M, & \lambda_{\max}^C \ge \lambda_{Total}^M; \\ \lambda_{\max}^C, & \lambda_{\max}^C < \lambda_{Total}^M. \end{cases}$$

边缘端平均等待时间:

$$T_{wait}^{C}\left(\lambda_{p}^{C}\right) = \frac{C\left(c, \rho^{C}\right)}{cu^{C} - \lambda_{p}^{C}} + \frac{1}{u^{C}}.$$

执行结果在边缘端的等待时间:

$$T_b^C \left(\lambda_p^C \right) = \frac{1}{u_b^C - \lambda_p^C}.$$

c: c个同类的边缘服务器 u^c : 每个边缘服务器的服务率 λ_{\max}^c : 边缘服务器的最大服务率 λ_{total}^M : MDs 占用的服务率

 $\psi^{\rm C}$: request在边缘端执行的概率

λ^C_p: 边缘服务器实际的服务率

uh: 边缘云的传输速率

(8)

(10) 包括排队时间和执行时间 (12)



论文模型: 中央端

• 中央云的排队模型为 M/M/∞

To: 中央端有线连接固定的延时

u^{cc}: 中央服务器的服务率

 λ_{total}^{M} : MDs 占用的服务率

λ^C_p: 边缘服务器实际的服务率

ubcc: 中央云的传输速率

超载的request等待时间:
$$T_{wait}^{CC} = T^O + \frac{1}{u^{CC}}$$
. (13)

执行结果在中央云端的等待时间: $T_b^{CC}\left(p_i^C\right) = \frac{1}{u_b^{CC} - \left(\lambda_{Total}^M - \lambda_p^C\right)}.$ (14)



论文模型

MD i 的平均能耗
$$E_i(p_i^C, P_i) = (1 - p_i^C) E_i^M(p_i^C, P_i) + p_i^C E_i^S(p_i^C, P_i)$$
. (15)

(17)

所有MD的平均能耗
$$E\left(p_i^C,P_i\right)=rac{1}{N}\sum_{i=1}^N E_i\left(p_i^C,P_i\right).$$

MD i 的平均执行时间
$$T_{i}\left(p_{i}^{C},P_{i}\right) = \left(1-p_{i}^{C}\right)T_{i}^{M}\left(p_{i}^{C}\right) + p_{i}^{C}T_{i}^{t}\left(p_{i}^{C},P_{i}\right) \\ + p_{i}^{C}\psi^{C}\left(T_{wait}^{C} + T_{b}^{C}\right) \qquad . \tag{16}$$

$$+ p_{i}^{C}\left(1-\psi^{C}\right)\left(T_{wait}^{CC} + T_{b}^{CC}\right)$$

所有MD的平均执行时间
$$T\left(p_i^C, P_i\right) = \frac{1}{N} \left[\sum_{i=1}^N T_i\left(p_i^C, P_i\right)\right].$$
 (18)



论文模型

所有MD的平均花费
$$M\left(p_{i}^{C}\right) = \frac{1}{N} \left\{r^{C} \lambda_{p}^{C}\left(p_{i}^{C}\right) + r^{CC}\left[\lambda_{Total}^{M}\left(p_{i}^{C}\right) - \lambda_{p}^{C}\left(p_{i}^{C}\right)\right]\right\}. \tag{19}$$

 r^c : 边缘服务器的服务单价

r^{cc}: 中央服务器的服务单价



论文模型

$$\min_{\left\{p_{i}^{C}, P_{i}\right\}} \left\{E\left(p_{i}^{C}, P_{i}\right), T\left(p_{i}^{C}, P_{i}\right), M\left(p_{i}^{C}\right)\right\}, \tag{20}$$

subject to

$$(1 - p_i^C) \lambda_i < u_i^M (1 - l_i^M), \qquad (21)$$

$$\lambda_p^C < cu^C, \tag{22}$$

$$\lambda_p^C < u_b^C, \tag{23}$$

$$\lambda_{Total}^{M} - \lambda_{p}^{C} < u_{b}^{CC}, \tag{24}$$

$$0 < P_i \le P^{th} \qquad (i = 1, 2, \dots, N),$$
 (25)

$$0 \le p_i^C \le 1 \qquad (i = 1, 2, \dots, N).$$
 (26)

谢谢

汇报人: 王琦

导师: 李智勇

