Throughput maximization for flows with fine input structure

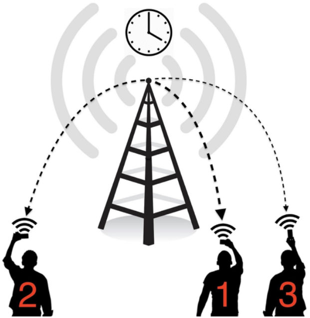
D. Pasechnyuk

Presidential Physics and Mathematics Lyceum №239, Saint-Petersburg

International Workshop "Optimization at work", 2017

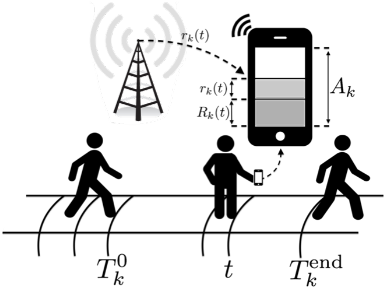
Problem Description

- **Cell tower** (base station) for mobile Internet traffic distribution.
- **Program-scheduler** for resource allocation among the base station clients.



Notation

- A_k — file size
- $R_k(t)$ — amount of transmitted information to the time t
- $r_k(t)$ — channel capacity in time t
- T_k^0 — time of the client's arrival in the system
- T_k^{end} — time of the client's departure from the system
- $\mathcal{Q}$ — queue clients



Quality Criteria

logALPT (Application Level Perceived Throughput)

$$logALPT = \lim_{T\to\infty} \frac{1}{N(T)} \sum_{k=1}^{N(T)} \log\left(\frac{A_k}{T_k^{end}(x) - T_k^0}\right) \to \max_x \qquad (1)$$

logALPT - the average logarithm of the actual speed of file transfer to clients

Index Strategy

Index Strategy

$$k^* = \arg\max_{k \in \mathcal{Q}} I(k) \tag{2}$$

Index I(k) is a measure of client's priority

Index Strategy

Examples

$$
\begin{aligned}
k^* &= \arg\max_{k\in\mathcal{Q}} \frac{r_k(t)}{R_k(t)/t} && \textbf{(PFS)} \\
k^* &= \arg\max_{k\in\mathcal{Q}} \frac{r_k(t)}{A_k - R_k(t)} && \textbf{(SRPT)} \\
k^* &= \arg\max_{k\in\mathcal{Q}} \frac{r_k(t)}{R_k(t)} && \textbf{(DAS)} \\
k^* &= \arg\max_{k\in\mathcal{Q}} \frac{1}{t - T_k^0} && \textbf{(PS)} \\
k^* &= \arg\max_{k\in\mathcal{Q}} r_k(t) && \textbf{(Max C/I)} \\
k^* &= \arg\max_{k\in\mathcal{Q}} \frac{r_k(t)}{t - T_k^0} && \textbf{(TAS)}
\end{aligned}
\qquad (3)
$$

Index Combinations

Linear combination

$$k^* = \arg\max_{k \in \mathcal{Q}} \sum_{i=1}^{n} \alpha_i I_i(k) \quad (4)$$

Probabilistic combination

$$k^* = \arg\max_{k \in \mathcal{Q}} \begin{cases} I_1(k) & \text{with probability } p_1 \\ I_2(k) & \text{with probability } p_2 \\ \vdots & \\ I_n(k) & \text{with probability } p_n \end{cases}, \sum_{i=1}^{n} p_i = 1 \quad (5)$$

Software

- Software for the simulation of emergence and processing of client's requests on the base station
- Programming language — Python 3
- Sources — https://github.com/pasechnyuk2004/Huawei

Parametric Model

- $T_k^{end} - T_k^0$ — random variables from exponential distribution

$$p(x) = Exp(x|\lambda) \quad (6)$$

$$Exp(x|\lambda) = \lambda e^{-\lambda x}, x > 0 \quad (7)$$

- A_k — random variables from a mixture of Pareto distributions

$$p(x) = \sum_{i=1}^{4} p_i \Pi(x|m_i, \alpha) \quad (8)$$

$$\Pi(x|m_i, \alpha) = \frac{\alpha m_i^{\alpha}}{x^{\alpha+1}}, x \geqslant m_i \quad (9)$$

- $r_k(t)$ — i.i.d random variables from uniform distribution

$$p(x) = U(x|a_k, b_k) \quad (10)$$

$$U(x|a_k, b_k) = \frac{1}{b_k - a_k}, x \in [a_k, b_k] \quad (11)$$

Results

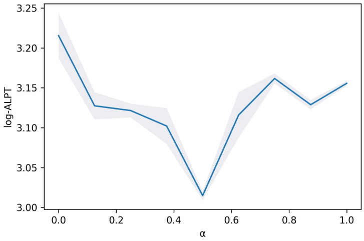


Figure: logALPT dependence on the parameter of the linear combination (DAS, TAS)

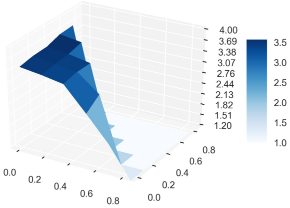


Figure: logALPT dependence on the parameters of the probabilistic combination (T, TAS, DAS)

$$\mathrm{I(k)} = \sum_{i=1}^{n} \alpha_i I_i(k)$$

$$\mathrm{I(k)} = \begin{cases} I_1(k) & \text{with prob. } p_1 \\ \vdots & \\ I_n(k) & \text{with prob. } p_n \end{cases}$$

Results

T & TK

$$
\begin{aligned}
k^* &= \arg\max_{k \in \mathcal{Q}} \frac{R_k(t)}{t - T_k^0 + \frac{R_k(t)}{\bar{r}_k(t)}} \qquad \textbf{(T)} \\
k^* &= \arg\max_{k \in \mathcal{Q}} \frac{r_k(t) R_k(t)}{t - T_k^0} \qquad \textbf{(TK)}
\end{aligned} \tag{12}
$$

Index	logALPT
T	4.01 ± 0.09
TK	3.93 ± 0.03
TAS	3.50 ± 0.05
Max C/I	1.38 ± 0.24
DAS	0.71 ± 0.14

(13)

Thank you
for your attention

Software configuration

- A_k:
  - $\alpha = 14.5$
  - $m_1 = 1000; p_1 = 0.4$
  - $m_1 = 10000; p_2 = 0.3$
  - $m_1 = 50000; p_3 = 0.2$
  - $m_1 = 125000; p_1 = 0.1$
- $r_k(t)$:
  - $a_k = 0.7\bar{r}$
  - $b_k = 1.3\bar{r}$
  - $\bar{r} = k\lambda\bar{A}_k$
  - $k = 0.8$
- *Simulation*:
  - initial clients count = 100
  - generated clients = 10000
  - total time of simulation = 258 simulated minutes
  - average queue length = 20 clients