Routing policies for periodic state information

Ernst Nordström

November 1, 2009

1 Introduction

The network is assumed to operate under on demand call set up, i.e call requests are lost when the network is found busy upon call arrival. The routing controller for each OD pair observes the network link states every time period τ seconds. Between observations calls are accepted and completed on different paths that results in random link state increments and decrements, respectively. Hence, the network state between two observations is uncertain. The latest state observation on link s is denoted $\mathbf{x}^s$. The corresponding decomposed network state is denoted $\mathbf{z} = \{\mathbf{x}^s\}$. At time $t \in [0, \tau)$ seconds after the last state observation the link s has state distribution $Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})$ (known as belief state).

The partially observable Markov decision process (POMDP) associated with the periodic state observations is non-Markovian. On the other hand, the belief MDP formed from the belief state distribution is Markov [1]. We refer only to the states between state updates via the belief state distribution. In particular we compute the exact time- and state-dependent randomized policy from the belief state distribution and the path shadow prices (obtained from the relative values) determined from the MDP with real-time state observations. We motivate the choice of the MDP with the *certainty equivalence principle* from stochastic control which allows the separation of control and estimation [3]. In other words the control problem with unknown realizations (from known probability distributions) can be formed by optimizing the control first under perfect forsight and then replacing the unknown future values by optimal forecasts.

The distribution $Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})$ can be obtained, for each link state $\mathbf{x}^s$, as solution to a differential equation system. The computations should be done at each policy iteration step using a numerical differential equation solver (e.g. Runge Kutta).

In this paper we outline five types of routing policies with increasing level of performance:

- static randomized policy,
- dynamic state-dependent deterministic policy,
- static state-dependent randomized policy,
- Approximate dynamic state-dependent randomized policy,
- Exact dynamic state-dependent randomized policy (optimal).

The state-dependent randomized routing policy $\pi(\mathbf{z}, t)$ require a large computational effort if applied on-line for each call arrival. A natural approach is to compute the routing policy off-line (once each policy iteration step). However, the the randomization weights must be computed for each combination of link states $\mathbf{z} = \{\mathbf{x}^s\}$ which makes the computational complexity prohibitive. To obtain a feasible solution of state-dependent randomized routing the probability that a given path has lower path shadow price than another path can not be computed in an exact manner. Also, note that the off-line computation requires discretization of time for the dynamic policy.

2 Traffic assumptions

The network is offered traffic from K classes which are, for sake of simplicity, subject to deterministic multiplexing. The j-th class, $j \in J = \{1, \ldots, K\}$, is characterized by the following:

- Origin-destination (OD) node pair,
- Bandwidth requirement b_j [Mbps],
- Poissonian call arrival process with rate λ_j [s^{-1}],
- Exponentially distributed call holding time with mean $1/\mu_j$ [s],
- Set of alternative routes, W_j,
- Reward parameter $r_j \in (0, \infty)$, and

The classes are classified into G bandwidth categories. The i-th category, $i \in I = \{1, \ldots, G\}$, is characterized by:

- Bandwidth requirement b_i [Mbps],
- Average mean call holding time $1/\overline{\mu}_i$ [s],
- Average reward parameter $\overline{r}_i$.

3 Belief state model

The transient distribution $Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})$ can be obtained, for each link state $\mathbf{x}^s$, as solution to a differential equation system. With the assumptions on exponential service time distribution and Poisson call arrivals we can write:

$$
\begin{aligned}
&Q^s(\mathbf{y}^s|\mathbf{x}^s, t + \Delta t, \mathbf{a}) = \\
&Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})[\textstyle\prod_{i\in I}(1 - y_i^s\overline{\mu}_i\Delta t)(1 - \lambda_i^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t) + \prod_{i\in I} y_i^s\overline{\mu}_i\Delta t\lambda_i^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t] \\
&+ \textstyle\sum_{i\in I} Q^s(\mathbf{y}^s - \delta_i|\mathbf{x}^s, t, \mathbf{a})[\lambda_i^s(\mathbf{y}^s - \delta_i|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t(1 - (y_i^s - 1)\overline{\mu}_i\Delta t)] \\
&+ \textstyle\sum_{i\in I} Q^s(\mathbf{y}^s + \delta_i|\mathbf{x}^s, t, \mathbf{a})[(y_i^s + 1)\overline{\mu}_i\Delta t(1 - \lambda_i^s(\mathbf{y}^s + \delta_i|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t)
\end{aligned}
$$

$$
\begin{aligned}
&Q^s(\mathbf{0}|\mathbf{x}^s, t + \Delta t, \mathbf{a}) = Q^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a}) \textstyle\sum_{i\in I}(1 - \lambda_i^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t) \\
&+ \textstyle\sum_{i\in I} Q^s(\delta_i|\mathbf{x}^s, t, \mathbf{a})[\overline{\mu}_i\Delta t(1 - \lambda_i^s(\delta_i|\mathbf{x}^s, t, \mathbf{a})a_i\Delta t)]
\end{aligned}
$$

where $\mathbf{0}$ denotes the empty state vector and δ_i denotes a vector with zeros expected for a one in position i. According to Taylor series we have

$$Q^s(\mathbf{y}^s|\mathbf{x}^s, t+\Delta t, \mathbf{a}) = Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})\frac{\mathrm{d}}{\mathrm{dt}}Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})\Delta t \tag{1}$$

Combining the two equations above and letting $\Delta t \to 0$ we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) = -\sum_{i\in I}(\lambda_i^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})a_i + y_i^s\overline{\mu}_i)Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) + \\ +\sum_{i\in I}Q^s(\mathbf{y}^s-\delta_i|\mathbf{x}^s, t, \mathbf{a})\lambda_i^s(\mathbf{y}^s-\delta_i|\mathbf{x}^s, t, \mathbf{a})a_i + \sum_{i\in I}Q^s(\mathbf{y}^s+\delta_i|\mathbf{x}^s, t, \mathbf{a})(y_i^s+1)\overline{\mu}_i\end{aligned} \tag{2}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}Q^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a}) = -\sum_{i\in I}\lambda_i^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a})a_iQ^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a}) + \sum_{i\in I}\overline{\mu}_iQ^s(\delta_i|\mathbf{x}^s, t, \mathbf{a}) \tag{3}$$

At $t = 0$ we know that the state is $\mathbf{x}^s$ with full certainty so we have the following initial condition:

$$\begin{aligned}&Q^s(\mathbf{x}^s|\mathbf{x}^s, 0, \mathbf{a}) = 1; \\ &Q^s(\mathbf{y}^s|\mathbf{x}^s, 0, \mathbf{a}) = 0, \qquad \mathbf{y}^s \in X^s \setminus \{\mathbf{x}^s\}\end{aligned} \tag{4}$$

The arrival rate in state $\mathbf{y}^s$ to link s, given link state $\mathbf{x}^s$ at the recent state update, is given by:

$$\lambda_i^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) = \lambda_i^s(\mathbf{x}^s, \pi)Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) \tag{5}$$

Hence, the system of differential equations can be written:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{dt}}Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) = -\sum_{i\in I}\lambda_i^s(\mathbf{x}^s, \pi)a_iQ^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})^2 - \sum_{i\in I}y_i^s\overline{\mu}_iQ^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) + \\ +\sum_{i\in I}Q^s(\mathbf{y}^s-\delta_i|\mathbf{x}^s, t, \mathbf{a})^2\lambda_i^s(\mathbf{x}^s, \pi)a_i + \sum_{i\in I}Q^s(\mathbf{y}^s+\delta_i|\mathbf{x}^s, t, \mathbf{a})(y_i^s+1)\overline{\mu}_i\end{aligned} \tag{6}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}Q^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a}) = -\sum_{i\in I}\lambda_i^s(\mathbf{x}^s, \pi)a_iQ^s(\mathbf{0}|\mathbf{x}^s, t, \mathbf{a})^2 + \sum_{i\in I}\overline{\mu}_iQ^s(\delta_i|\mathbf{x}^s, t, \mathbf{a}) \tag{7}$$

This is a non-linear differential equation system which can be solved by some numerical method, e.g. the Runge-Kutta method.

4 Static randomized policy

The static routing policy π for a class j call request is specified by the constant randomization weights h_j^k:

$$h_j^k = Pr\left\{\mathbf{p}^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}, \mathbf{p}^k < r_j\right\} \tag{8}$$

where $\mathbf{p}^k$ denote the random path shadow price for path k. The call request is rejected with probability $h_j^0 = 1 - \sum_{k \in W_j} h_j^k$. The weights can be written:

$$h_j^k = \sum_{\mathbf{x}^1 \in X^1} \cdots \sum_{\mathbf{x}^n \in X^{n_k}} Pr\left\{p^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}, \mathbf{p}^k < r_j\right\} \prod_{s=1}^{n_k} Q^s(\mathbf{x}^s) \tag{9}$$

where path $k \in W_j$ consist of n_k links. Since paths are independent by assumption we have

$$h_j^k = \sum_{\mathbf{x}^1 \in X^1} \cdots \sum_{\mathbf{x}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} Pr\left\{p^k < \mathbf{p}^l\right\} \theta(r_j - p^k) \prod_{s=1}^{n_k} Q^s(\mathbf{x}^s) \tag{10}$$

where $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x \leq 0$ and p^k denotes the deterministic shadow price for path k in state $(\mathbf{x}^1, \ldots, \mathbf{x}^{n_k})$. Let

$$F_{\mathbf{p}}^k(v) = Pr\{\mathbf{p}^k \leq v\} \tag{11}$$

Define

$$G_{\mathbf{p}}^k(v) = 1 - F_{\mathbf{p}}^k(v) \tag{12}$$

By definition of $G_{\mathbf{p}}^k(v)$ we have

$$Pr\{\mathbf{p}^l > p^k\} = G_{\mathbf{p}}^l(p^k) \tag{13}$$

Which gives the result

$$h_j^k = \sum_{\mathbf{x}^1 \in X^1} \cdots \sum_{\mathbf{x}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} G_{\mathbf{p}}^l(p^k)\theta(r_j - p^k) \prod_{s=1}^{n_k} Q^s(\mathbf{x}^s) \tag{14}$$

The probability $G_{\mathbf{p}}^l(p^k)$ can be obtained as:

$$G_{\mathbf{p}}^l(p^k) = \sum_{\mathbf{x}^1 \in X^1} \cdots \sum_{\mathbf{x}^{n_l} \in X^{n_l}} \theta(p^l(\mathbf{x}^1, \ldots, \mathbf{x}^{n_l}) - p^k) \prod_{s=1}^{n_l} Q^s(\mathbf{x}^s) \tag{15}$$

5 Dynamic deterministic policy

A simple deterministic policy with relatively low complexity is as follows. The new call is then allocated to the path, among the set of feasible paths, with the largest positive average path net-gain $\overline{g}_j^k(\mathbf{z}, t, \pi)$:

$$\overline{g}_j^k(\mathbf{z}, t, \pi) = r_j - \sum_{s \in S_k} \overline{p}_i^s(\mathbf{x}^s, t, \pi), \tag{16}$$

The average link shadow prices $\overline{p}_i^s(\mathbf{x}^s, t, \pi)$ is given by:

$$\overline{p}_i^s(\mathbf{x}^s, t, \pi) = \overline{r}_i^s(\pi) - \overline{g}_i^s(\mathbf{x}^s, t, \pi), \tag{17}$$

where the average link net-gain $\overline{g}_i^s(\mathbf{x}^s, t, \pi)$ is obtained from the belief state distribution $Q^s(\mathbf{y}^s|\mathbf{x}^s, t)$:

$$\overline{g}_i^s(\mathbf{x}^s, t, \pi) = \sum_{\mathbf{y}^s \in X} Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) g_i^s(\mathbf{y}^s, \pi) \tag{18}$$

and from the link net-gain $g_i^s(\mathbf{y}^s, \pi)$:

$$g_i^s(\mathbf{y}^s, \pi) = v^s(\mathbf{y}^s + \delta_i, \pi) - v^s(\mathbf{y}^s, \pi), \tag{19}$$

where $v^s(\mathbf{y}^s, \pi)$ denotes the relative value in state $\mathbf{y}^s$ of link s.

6 Static state-dependent randomized policy

The static state-dependent routing policy $\pi(\mathbf{z})$ for a class j call request is specified by the randomization weights $h_j^k(\mathbf{z})$:

$$h_j^k(\mathbf{z}) = Pr\left\{\mathbf{p}^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}, \mathbf{p}^k < r_j\right\} \tag{20}$$

where $\mathbf{p}^k$ denote the random path shadow price for path k. The call request is rejected with probability $h_j^0(\mathbf{z}) = 1 - \sum_{k \in W_j} h_j^k(\mathbf{z})$. The weights can be written:

$$h_j^k(\mathbf{z}) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} Pr\left\{p^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}\right\} \theta(r_j - p^k) \prod_{s=1}^{n_k} \frac{1}{\tau} \int_0^\tau Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) dt \tag{21}$$

where path $k \in W_j$ consist of n_k links and $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x \leq 0$. Since paths are independent by assumption we have

$$h_j^k(\mathbf{z}) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} Pr\left\{p^k < \mathbf{p}^l\right\} \theta(r_j - p^k) \prod_{s=1}^{n_k} \frac{1}{\tau} \int_0^\tau Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) dt \tag{22}$$

where p^k denotes the deterministic shadow price for path k in state $(\mathbf{y}^1, \ldots, \mathbf{y}^{n_k})$. Let

$$F_{\mathbf{p}}^k(v) = Pr\{\mathbf{p}^k \leq v\} \tag{23}$$

Define

$$G_{\mathbf{p}}^k(v) = 1 - F_{\mathbf{p}}^k(v) \tag{24}$$

By definition of $G_{\mathbf{p}}^k(v)$ we have

$$Pr\{\mathbf{p}^l > p^k\} = G_{\mathbf{p}}^l(p^k) \tag{25}$$

Which gives the result

$$h_j^k(\mathbf{z}) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} G_{\mathbf{p}}^l(p^k)\theta(r_j - p^k) \prod_{s=1}^{n_k} \frac{1}{\tau} \int_0^\tau Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})dt \tag{26}$$

The probability $G_{\mathbf{p}}^l(p^k)$ can be obtained as:

$$G_{\mathbf{p}}^l(p^k) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^{n_l} \in X^{n_l}} \theta(p^l(\mathbf{y}^1, \ldots, \mathbf{y}^{n_l}) - p^k) \prod_{s=1}^{n_l} \frac{1}{\tau} \int_0^\tau Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a})dt \tag{27}$$

The probability $G_{\mathbf{p}}^l(p^k)$ consitutes a major computational burden.

7 Exact dynamic state-dependent randomized policy

The time- and state-dependent routing policy $\pi(\mathbf{z}, t)$ for a class j call request is specified by the randomization weights $h_j^k(\mathbf{z}, t)$:

$$h_j^k(\mathbf{z}, t) = Pr\left\{\mathbf{p}^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}, \mathbf{p}^k < r_j\right\} \tag{28}$$

where $\mathbf{p}^k$ denote the random path shadow price for path k. The call request is rejected with probability $h_j^0(\mathbf{z}, t) = 1 - \sum_{k \in W_j} h_j^k(\mathbf{z}, t)$. The weights can be written:

$$h_j^k(\mathbf{z}, t) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} Pr\left\{p^k < \mathbf{p}^l, \forall l \in W_j \setminus \{k\}\right\} \theta(r_j - p^k) \prod_{s=1}^{n_k} Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) \tag{29}$$

where path $k \in W_j$ consist of n_k links and $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x \le 0$. Since paths are independent by assumption we have

$$h_j^k(\mathbf{z}, t) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} Pr\left\{p^k < \mathbf{p}^l\right\} \theta(r_j - p^k) \prod_{s=1}^{n_k} Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) \tag{30}$$

where p^k denotes the deterministic shadow price for path k in state $(\mathbf{y}^1, \ldots, \mathbf{y}^{n_k})$. Let

$$F_{\mathbf{p}}^k(v) = Pr\{\mathbf{p}^k \le v\} \tag{31}$$

Define

$$G_{\mathbf{p}}^k(v) = 1 - F_{\mathbf{p}}^k(v) \tag{32}$$

By definition of $G_{\mathbf{p}}^k(v)$ we have

$$Pr\{\mathbf{p}^l > p^k\} = G_{\mathbf{p}}^l(p^k) \tag{33}$$

Which gives the result

$$h_j^k(\mathbf{z}, t) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^n \in X^{n_k}} \prod_{l \in W_j \setminus \{k\}} G_{\mathbf{p}}^l(p^k)\theta(r_j - p^k) \prod_{s=1}^{n_k} Q^s(\mathbf{y}^s|\mathbf{x}^s, t, \mathbf{a}) \tag{34}$$

The probability $G^l_{\mathbf{p}}(p^k)$ can be obtained as:

$$G^l_{\mathbf{p}}(p^k) = \sum_{\mathbf{y}^1 \in X^1} \cdots \sum_{\mathbf{y}^{n_l} \in X^{n_l}} \theta(p^l(\mathbf{y}^1, \ldots, \mathbf{y}^{n_l}) - p^k) \prod_{s=1}^{n_l} Q^s(\mathbf{y}^s | \mathbf{x}^s, t, \mathbf{a}) \tag{35}$$

8 Approximate dynamic state-dependent randomized policy

For every state of path k the probability $G^l_{\mathbf{p}}(p^k)$ that path k has lower path shadow price than the competing path l must be computed The exact computation requires $O(S^{n_l})$ operations where S denotes the maximum size of the state space of the links in path l.

One solution attempt is to represent each link state on the competing path l by the equilibrium distribution $Q^s(\mathbf{x}^s)$:

$$G^l_{\mathbf{p}}(p^k) = \sum_{\mathbf{x}^1 \in X^1} \cdots \sum_{\mathbf{x}^{n_l} \in X^{n_l}} \theta(p^l(\mathbf{x}^1, \ldots, \mathbf{x}^{n_l}) - p^k) \prod_{s=1}^{n_l} Q^s(\mathbf{x}^s) \tag{36}$$

In this case the randomization weight for path k only depend on the state $\mathbf{z}$ of this path, and not on the states of the competing paths.

References

[1] Aström K.J., "Optimal control of Markov decision processes with incomplete state information",*Journal of Mathematical Analysis and Applications*, 10:174-205, 1965.

[2] Dziong Z., Mignault J., Rosenberg C., "Blocking evaluation for networks with reward maximization routing", In *Proceedings of INFOCOM'93*, pp. 593-601, San Fransisco, USA, 1993.

[3] Söderström T., *Discrete-time Stochastic Systems–Estimation and Control*, 2nd edition, Springer-Verlag, 2002.