A First Step Towards the Resolution of The Starvation Problem In Multi-Point-to-Point ICRCNs

Wissam Fawaz, Ribal Atallah and Maurice Khabbaz

***Abstract*—This letter revolves around an Intermittently Connected Roadside Communication Network (ICRCN) scenario consisting of isolated source Stationary Roadside Units (SRUs) exploiting mobile smart vehicles as store-carry-forward data relays to a destination SRU. In this case, it is observed that a subset of these source SRUs may suffer from a significant starvation problem. In this letter, first, an Markov Decision Process (MDP) framework is established for the purpose of identifying a suitable Bulk Release Decision Policy (BRDP). Second, BRDP is implemented within a Starvation Mitigation and Delay-Minimal (SMDM) bundle delivery scheme. Extensive simulations are conducted for the purpose of: *a*) quantifying the severity of the starvation experienced by the downstream SRUs and *b*) gauging the merit of the proposed SMDM scheme through its ability to jointly mitigate starvation and achieve end-to-end delay minimal bundle delivery to the destination SRU.**

***Index Terms*—ICRCN, SRU, Performance Evaluation, MDP.**

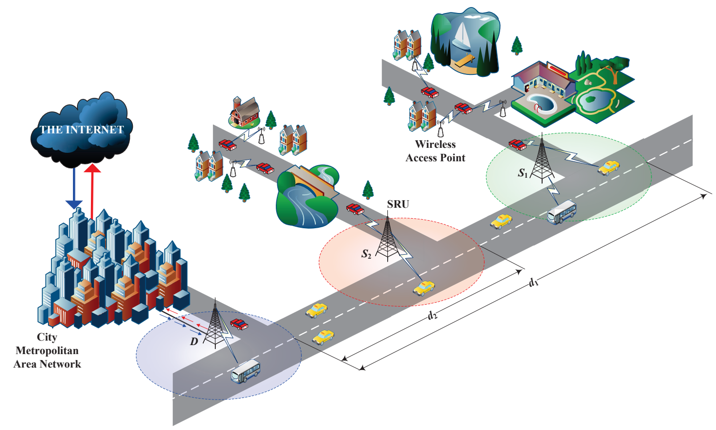


Fig. 1. ICRCN sub-network scenario.

I. Introduction

The utilization of the transportation infrastructure as a means for establishing connectivity among isolated Stationary Roadside Units (SRUs) represents an emerging terrestrial application of the Disruption-Tolerant Networking (DTN) paradigm, [1]. Recently, this application has gained significant momentum. This is especially true since it has been proven to be an effective and cost-minimal solution for bringing digital connectivity to rural areas, where the cost of setting up a networking infrastructure can be elevated [2]. This letter considers a networking scenario such as the one depicted in Figure 1 which consists of three SRUs, namely: a) S_1 and S_2 being two source SRUs and b) D being a destination SRU. All three SRUs are deployed along a one-dimensional and uninterrupted roadway segment. Each one of these three SRUs is located outside the respective coverage ranges of the two others and hence the three SRUs cannot directly communicate with one another. Furthermore, only D is connected to the Internet through minimal networking infrastructure. The two source SRUs S_1 and S_2 are completely isolated. In the absence of networking infrastructure connecting S_1 and S_2 to D, mobile vehicles equipped with computerized modules, finite buffers and wireless communication devices serve as store-carry-forward data carriers from both of S_1 and S_2 to D. Such a networking scenario belongs to the class of Intermittently Connected Roadside Communication Networks (ICRCNs). Upon entering the communication range[1] of either one of the two source SRUs S_1 or S_2, a vehicle presents to that SRU a bundle[2] release opportunity. In the context of a point-to-point ICRCN with a single source SRU, the work of [2] aimed at determining the suitability of such an arising opportunity in terms of the minimization of the average end-to-end delivery delay of a singly released bundle. In a similar context, the authors of [4] addressed the limitations of [2] and demonstrated how the release of a subset of the bundles[3] queueing in a source SRU's buffer results in a significant performance improvement. Throughout this present work, it is observed that for the duration of their journey over the roadway segment illustrated in Figure 1, vehicles will first enter the coverage range of S_1 and then, on their way to D, they will pass by S_2. In turn, each of S_1 and S_2 will opportunistically attempt to load the arriving vehicles with as many as possible of their respective data bundles, if available. However, since the vehicles' buffer sizes are finite, the middle source SRU, S_2, is highly likely to suffer from a bundle release restriction or even a denial of bundle release. This is particularly true since the buffer of an arriving vehicle to S_2 might have been considerably loaded (ultimately fully exhausted) with S_1's bundles. Typically, a node such as S_2 is referred to as a starving node since bundles would rapidly accumulate in S_2's buffer and thus would forcefully experience excessive queueing delays. What distinguishes this letter from the work of [2], [4] is that it aims at both highlighting and mitigating the severity of the starvation that a downstream SRU such as S_2 may suffer from.

W. Fawaz and R. Atallah are with the ECE department of the Lebanese American University. E-Mail: {*wissam.fawaz*, *ribal.atallah*}@lau.edu.lb

M. Khabbaz is with the ECE department of the Notre-Dame University. E-Mail: *maurice.khabbaz*@gmail.com

[1] In the sequel, the event of a vehicle entering the communication range of an SRU is referred to as a vehicle arrival to that SRU.

[2] Data and control signals are combined in a single atomic entity, called bundle, that is transmitted across a DTN-based ICRCN, [1].

[3] Thereafter, a subset/group of bundles is referred to as a bulk of bundles.

II. Identification of the Starvation Problem

In the context of the ICRCN sub-network illustrated in Figure 1, both S_1 and S_2 are assumed to be identical. Their respective coverage ranges span non-overlapping and identical road segments of length d_C. The distances that separate S_1 and S_2 from D are d_1 and d_2 respectively. An arbitrary vehicle k navigating at a speed v_k over the roadway segment $[S_1D]$ arrives to S_1 at time $t_{1,k}$. It resides within S_1's communication range for a period of time $R_k = \frac{d_C}{v_k}$ known as the vehicle's residence time. As such, vehicle k presents an opportunity for S_1 to release a bulk of the bundles queueing in its buffer. In turn, vehicle k will carry this bulk to D. While steering towards D, vehicle k encounters S_2. This latter also attempts to release a bulk to vehicle k. However, observe that S_2 may not be able to take advantage of vehicle k's entire residence time to download bundles to that vehicle. An in-house JAVA-based discrete-event simulator is developed to conduct extensive simulations in order to validate this observation. Throughout a subset of the conducted simulations, the two source SRUs S_1 and S_2 were set to operate under the Greedy Bundle Release Scheme with Bulk Bundle Release (GBRS-BBR) which was developed in [4]. Under GBRS-BBR, a source SRU greedily releases a bulk of bundles to every arriving vehicle irrespective of that vehicle's speed. Next, these simulations are repeated while setting S_1 and S_2 to operate under a modified version of the scheme[4] proposed in [2] referred to hereafter as the Delay-Optimal Bulk Delivery (DOBD). Under DOBD, a source SRU is instructed to release a bulk of bundles only to the vehicle that contributes the most to the minimization of the average end-to-end delivery delay. In addition, the following assumptions were borrowed from [2], [3], [4]:

- *A*1: Bundle interarrival times are exponentially distributed with a probability density function $f_B(t) = \lambda_b e^{-\lambda_b t}$, where $t \geq 0$.
- *A*2: S_1 and S_2 are subject to equal bundle arrival rates of λ_b.
- *A*3: The size of each bundle is constant and denoted by b.
- *A*4: S_1 and S_2 have equal data transmission rates of T_R.
- *A*5: S_1 and S_2 are equipped with buffers having infinite sizes.
- *A*6: The vehicles arrive according to a Poisson process with a rate of μ_v.
- *A*7: The per-vehicle speed V_k is uniformly distributed in the range $[V_{min}; V_{max}]$.
- *A*8: The per-vehicle capacity C_k is uniformly distributed in the range $[C_{min}; C_{max}]$.
- *A*9: Each vehicle maintains a constant speed during its navigation over the roadway segment $[S_1D]$.
- *A*10: The IEEE 802.11p protocol is used for vehicle-to-SRU communication.

The adopted simulation input parameter values are as follows: *a)* $\lambda_b = 30$ (bundles/s), *b)* $b = 1500$ (bytes), *c)* $T_R = 1$ (Mbits/s), *d)* $\mu_v \in [0.004; 0.033]$ (vehicles/s), *e)* $V_{min} = 10$ (m/s) *f)* $V_{max} = 50$ (m/s), *g)* $C_{min} = 0$, *h)* $C_{max} = 30000$, *i)* $d_C = 200$ (m), *j)* $d_1 = 2000$ (m) and *k)* $d_2 = 1000$ (m). At this level, it is worthwhile mentioning that ρ, the overall load of the system composed of both S_1 and S_2, is:

$$\rho = \frac{(\lambda_1 + \lambda_2)}{\mu_v \overline{C}} = \frac{2\lambda_b}{\mu_v \overline{C}} \tag{1}$$

where λ_i is the bundle arrival rate to the i^{th} SRU ($i = 1, 2$) and $\overline{C}$ is the average vehicle storage capacity. The above-enumerated simulation input parameter values were chosen in such a way that the resulting system of SRUs is stable with $0 < \rho \leq 1$.

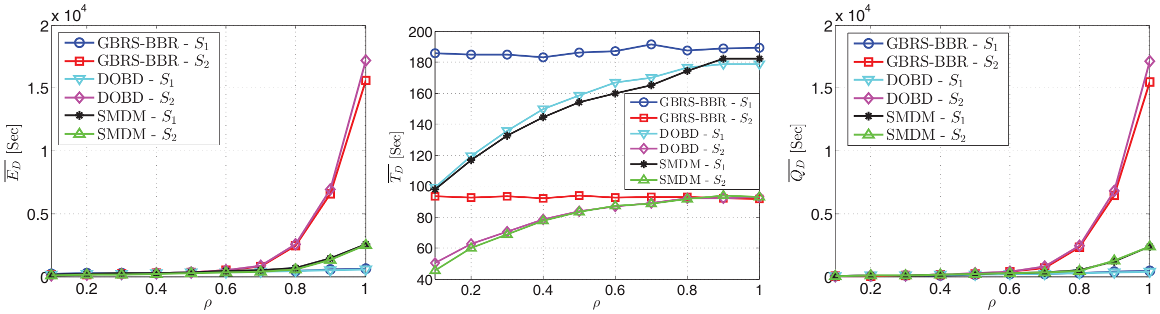
Figure 2(a) plots the average bundle end-to-end delivery delay, denoted by $\overline{E_D}$, achieved by GBRS-BBR and DOBD from both of the source SRUs S_1 and S_2 as a function of ρ. The figure clearly indicates that, as opposed to the bundles of S_1, those of S_2 experience remarkably excessive end-to-end delays. This finding is better explained if $\overline{E_D}$ is decomposed into its two underlying components, namely: *a*) the average bundle transit delay $\overline{T_D}$ and *b*) the average bundle queueing delay $\overline{Q_D}$. As illustrated in Figure 2(b), the bundles released at S_2 experience shorter transit delays as compared to those generated by S_1. This is simply justified by the fact that, before being dumped at the destination SRU D, the bundles produced by S_2 travel a shorter distance than that traveled by S_1's bundles. However, as far as S_2 is concerned, what is relevant to the present study is the fact that, in comparison to $\overline{Q_D}$ in Figure 2(c), $\overline{T_D}$ is almost two orders of magnitude smaller. As such, it becomes obvious that $\overline{Q_D}$ experienced by S_2's bundles is the bottleneck. Indeed, Figures 2(a) and 2(c) are clear evidence that $\overline{Q_D}$ overshadows $\overline{T_D}$ and governs the behavior of $\overline{E_D}$. The large $\overline{Q_D}$ experienced by S_2 is the result of a rapid accumulation of bundles in the buffer of that latter. As such, S_2 is said to suffer from a starvation problem.

III. A Novel Bulk Release Decision Policy

A. Theoretical Analysis

Upon the occurrence of a bulk release opportunity (*i.e.* a vehicle enters the communication range of a source SRU otherwise referred to as a vehicle arrival to the source SRU), either one of S_1 or S_2 has to make a decision as to whether or not to exploit that arriving vehicle as a bulk carrier to the destination SRU D. After the arrival of the first vehicle, the release decision process evolves over the sequence of vehicles $\{0, 1, 2, ..., k, ...\}$ with speeds $\{v_0, v_1, v_2, ..., v_k, ...\}$ arriving at times $\{t_0, t_1, t_2, ..., t_k, ...\}$. Following the assumption (*A*6) in section II, $\{I_k, k \geq 1\}$ forms a sequence of independent and identical exponentially distributed vehicle inter-arrival time intervals with an average $E[I] = \frac{1}{\mu_v}$. Denote by $N_{i,k}$ the number of bundles queueing in the i^{th} source SRU's buffer ($i = 1, 2$) at time t_k. Also let $Y_{i,k}$ represent the decision made by the i^{th} SRU ($i = 1, 2$) at time t_k and $X_{i,k}$ represent the size of the bulk released by the i^{th} source SRU at that time. In the context of the ICRCN sub-network scenario illustrated in Figure 1, $Y_{i,k} \in \{0; 1\}$. Observe, however, that, whenever

[4]The scheme proposed in [2] enables an SRU to release a single bundle per opportunity. For consistency with GBRS-BBR herein, this scheme was modified to enable the release of a bulk of bundles per opportunity.



(a) Average end-to-end delay. (b) Average bundle transit delay. (c) Average bundle queueing delay.

Fig. 2. Delay vs. load under the GBRS-BBR, DOBD, and SMDM schemes.

$N_{i,k} = 0$, $Y_{i,k} = 0$ is the only possible decision to be made and thus no bundles are released. As such:

$$N_{i,k+1} = max\{0, N_{i,k} - X_{i,k} \cdot Y_{i,k}\} + A_{i,k} \tag{2}$$

where $A_{i,k}$ represents the number of newly incoming bundles during the time interval $I_{k+1} = t_{k+1} - t_k$. $A_{i,k}$ has an average $\overline{A_i} = E[A_{i,k}] = \lambda_b E[I]$. In light of the above, the state of the middle source SRU S_2 can be described by the two-tuple $(N_{2,k}; V_k)$ and its dynamics are governed by equation (2). Now, define:

$$C_2\left(\ddot{n}_k, v_k, \ddot{y}_k\right) = \frac{\ddot{n}_k}{A_2} + f_2(v_k, \ddot{y}_k) \tag{3}$$

$C_2\left(\ddot{n}_k, v_k, \ddot{y}_k\right)$ is interpreted as the single stage cost function associated with S_2's state $(N_{2,k} = \ddot{n}_k; V_k = v_k)$ and release decision $Y_{2,k} = \ddot{y}_k$. It is composed of two terms, namely: *a*) $\frac{\ddot{n}_k}{A_2}$ being the cost associated with S_2's queue length (or the queueing delay) and *b*) $f_2(v_k, \ddot{y}_k)$ corresponding to the cost incurred by the exploitation of the vehicle whose speed is v_k as a bundle transporter from S_2 to D (i.e., the transit delay). Note that $f_2(v_k, \ddot{y}_k) = \frac{d_2}{A_2} \cdot \frac{\ddot{y}_k}{v_k}$. Given that our objective is to jointly minimize the queueing delay and the transit delay of bundles at S_2, it follows that the single stage cost function given in (3) establishes the basis for deriving an optimal average delay minimization policy $\pi_2^* \equiv \pi_2^*(\ddot{n}_k, v_k, \ddot{y}_k)$ at S_2.

At this level, it follows from assumptions ($A6$) and ($A7$) in section II that, first, V_k is uniformly distributed over a finite range and, second, the vehicle inter-arrival time is exponentially distributed. In addition, the decision space $Y_{2,k}$ is finite. As such, based on the knowledge acquired from [2], there exists a policy π_2^* that minimizes the following cost function at S_2:

$$\overline{C_2} = \frac{1}{A_2} \limsup_{k\to\infty} \left[\frac{1}{k} \cdot E\left[\sum_{i=0}^{k}\left(N_{2,i} + \frac{d_2}{V_i} \cdot Y_{2,i}\right)\right]\right] \tag{4}$$

In [5], the authors looked into a somewhat similar optimization problem in the context of a single downlink wireless channel. In particular, they considered the case of a slotted wireless channel where the transmission rate in a slot i depended on the allocated power P_i. Packets were assumed to arrive at an infinite buffer and the objective was to minimize both the queueing delay experienced by the packets and the average power required to support the arrival process. The single cost function used in [5] was as follows: $X_i + \beta P_i$; with X_i being the queue length in the i^{th} slot, P_i the power allocated during that slot, and β a Lagrangian multiplier. Their optimization problem boiled down to identifying a policy π that minimizes the following objective function:

$$\limsup_{k\to\infty} \left[\frac{1}{k} \cdot E\left[\sum_{i=0}^{k}\left(\frac{X_i}{\lambda} + \beta \cdot P_i\right)\right]\right]$$

So, it is clear that our problem formulation is very much inspired by the formulation in [5]. Hence, the results and findings reported in [5] with respect to the derivation of the optimal policy can be directly extended to our framework after: a) mapping the average power in [5] to the average transit delay in our study and b) setting the value of β to 1. More formally, the problem of finding the policy π_2^* that minimizes (4) is a classical average cost Markov decision problem where at each time step, S_2 chooses an action, namely, a release/no-release decision, and incurs a per-stage cost analogous to the one given by (3). Such problems can be solved via dynamic programming techniques such as the ones discussed in [6].

At this point and similarly to S_2, a single stage cost function needs to be formulated for S_1. Doing this is not trivial since, in addition to S_1's queue size and release decision, this cost function has to account for S_2's queue size and release decision as well. These two pieces of information can for instance be acquired by S_1 in a non-real time manner through the vehicles navigating from S_2 to S_1. More sophisticated estimation techniques can be used to this end. These, however, are beyond the scope of the present study and are left for future work. The cost function is, therefore, expressed as:

$$C_1(\dot{n}_k, \ddot{n}_k, v_k, \dot{y}_k, \ddot{y}_k) = \frac{\dot{n}_k}{A_1} + f_1(v_k, \dot{y}_k) + C_2\left(\ddot{n}_k, v_k, \ddot{y}_k\right) \tag{5}$$

Similar to (3), the term $\frac{\dot{n}_k}{A_1}$ in (5) is tied to the queue length of S_1. Also, in the same spirit, $f_1(v_k, \dot{y}_k) = \frac{d_1}{A_1} \cdot \frac{\dot{y}_k}{v_k}$. Moreover, the integration of S_2's cost function into that of S_1 is equivalent to feeding S_1 with enough information regarding the status of S_2's buffer as well as this latter's upcoming release decision. This obviously leads to properly adjusting S_1's decision as to whether or not to exploit the arriving vehicle whose speed is v_k as a bundle transporter to D. Following similar arguments as above, there exists a deterministic Markov policy $\pi_1^* \equiv \pi_1^*(\dot{n}_k, \ddot{n}_k, v_k, \dot{y}_k, \ddot{y}_k)$ that minimizes the average cost function developed for S_1 and which is given by:

$$\overline{C_1} = \frac{1}{\overline{A_1}} \limsup_{k\to\infty} \left[\frac{1}{k} E \left[\sum_{i=0}^{k} N_{1,i} + \frac{d_1}{V_i} \cdot Y_{1,i} \right] \right] + \overline{C_2} \quad (6)$$

To this end, the joint starvation mitigation/delay minimization problem at S_1 has been formulated into a Markov Decision Process framework with an average cost criterion given by (6). Following the guidelines of [6] and an analysis similar to the one presented for S_2, this problem may be resolved using a Dynamic Programming approach. Having identified the suitable policies π_1^* and π_2^* respectively for S_1 and S_2, the Bulk Release Decision Policy (BRDP) for the overall system of the two source SRUs is represented by $\pi^* \equiv (\pi_1^*, \pi_2^*)$. Next, BRDP is implemented within a Starvation Mitigation Delay-Minimal (SMDM) bulk bundle delivery scheme.

B. SMDM Simulations

Simulations are conducted using the simulator developed in section II for the purpose of evaluating the performance of the SMDM scheme. In the context of the SMDM-related simulation study, the bundle release decisions made by both source SRUs S_1 and S_2 are driven mainly by π^*. In addition, if a decision is made at time t_k to release a bulk of bundles to an arbitrary vehicle k with a free buffer capacity C_k, the number of bundles $X_{i,k}$ to be downloaded to that vehicle by the i^{th} SRU $(i = 1, 2)$ is determined as follows. At S_1, $X_{1,k}$ basically depends on: *a*) C_k being the amount of free buffer space that is available at the k^{th} vehicle as well as on *b*) $N_{2,k}$ being the number of bundles currently queueing at S_2. More specifically, the number of bundles released by S_1 to a vehicle k selected by BRDP is governed by the following inequality:

$$X_{1,k} \leq \frac{N_{1,k}}{N_{1,k} + N_{2,k}} \cdot C_k \quad (7)$$

In this way, $N_{2,k}$ plays a major role in throttling S_1 to which, now, is allocated only a well defined portion of the free buffer space available at the vehicle chosen by BRDP. This, by itself, contributes to the reservation of some spare buffer space for S_2. Consequently, S_2 will, in turn, beneficially exploit the more abundant remaining vehicle buffer capacity spared by S_1 and hence abandon the title of *a starving node*. Figure 2(a) compares the average end-to-end bundle delivery delays obtained under GBRS-BBR and DOBD to those realized by SMDM. This figure reflects the remarkable ability of the SMDM scheme in reducing the average end-to-end bundle delivery delay experienced by S_2's bundles. Knowing that both S_1 and S_2 are set to operate under the exact same conditions as those considered for the simulated scenarios in section II, the only reason for this significant performance improvement is the fact that SMDM enables S_2 to take further advantage of the arising bulk release opportunities, and hence clear out bulks of larger sizes. This implies the resolution of the starvation problem suffered by S_2. Figure 2(b) concurrently plots the transit delay from both S_1 and S_2 as achieved by all of the GBRS-BBR, DOBD and SMDM schemes. Under GBRS-BBR, a source SRU releases bundles to all of the arriving vehicles irrespective of their speeds. In other words, all of the slow and fast vehicles will receive bundles from the source SRUs. Consequently, the average transit delay achieved by GBRS-BBR is constant irrespective of the load on both source SRUs. Furthermore, note that when the network load approaches unity, the average transit delay achieved under either one of the DOBD or the SMDM schemes converges to that achieved under the GBRS-BBR scheme. This is due to the fact that, whenever the network load increases, bundles arrive to source SRUs in an accelerated trend. As a result the release policy instructs the SRUs to attempt a faster bundle clearance. Given that the only means for faster bundle clearance is to release them more frequently to arriving vehicles, the SRUs become obliged to select yet slower vehicles as relays in order to compensate for the rapid bundle accumulation in the queue. Eventually, the source SRUs will exploit all passing by vehicles as store-carry-forward relays, and thus the average transit delay of bundles will converge to that achieved under the GBRS-BBR scheme. Figure 2(c) concurrently plots the curves corresponding to the average queuing delay achieved respectively at S_1 and S_2 under all of the GBRS-BBR, DOBD and SMDM schemes. This figure reflects the remarkable ability of SMDM to resolve the starvation problem suffered by the middle source SRU S_2. When a bundle release opportunity arises, S_1 follows the BRDP implemented within SMDM and limits the amount of bundles it releases to the arriving vehicle. As such, S_1 may no longer selfishly exhaust the storage capacity of the arriving vehicles. Instead, it leaves room for S_2 to clear out as many bundles as possible to avoid starvation.

IV. Conclusion

This letter discussed a first step towards the resolution of a starvation problem that may arise in the context of emerging ICRCNs. An MDP framework is developed for the purpose of identifying a bulk release decision policy (BRDP) suitable for reducing/mitigating this problem at interposed SRUs. BRDP is implemented within a Starvation Mitigation Delay-Minimal (SMDM) scheme which is deployed at all SRUs. A simulation study revealed the ability of SMDM to resolve the starvation problem while achieving delay-minimal bulk delivery to the destination SRU.

References

[1] V. Cerf (*et al.*), "Delay-Tolerant Network Architecture", *IETF*, RFC 4838, April 2007.
[2] V. Ramaiyan (*et al.*), "Delay-Optimal Schedule for a Two-Hop Vehicular Relay Network", *ACM/Springer Mobile Networks and Applications*, 15:1, May 2009.
[3] M. Khabbaz (*et al.*), "Probabilistic Bundle Relaying Schemes in Two-Hop Vehicular Delay Tolerant Networks", *IEEE Communications Letters*, 15:3, March 2011.
[4] M. Khabbaz (*et al.*), "A Probabilistic and Traffic-aware Bundle Release Scheme for Vehicular Intermittently Connected Networks", *IEEE Transactions on Communications*, 60:11, November 2012.
[5] R.A. Berry (*et al.*), "Communication over Fading Channels with Delay Constraints", *IEEE Transactions on Information Theory*, 48:5, May 2002.
[6] D. Bertsekas "Dynamic Programming and Optimal Control", *Athena Scientific*, 2012, vol. I and II.