A Markovian ROHC Control Mechanism Based on Transport Block Link Model in LTE Networks

Carlos Feres
Electrical and Computer Engineering
University of California, Davis
Davis, CA, US
cjferes@ucdavis.edu

Wenhao Wu
Electrical and Computer Engineering
University of California, Davis
Davis, CA, US
wnhwu@ucdavis.edu

Zhi Ding
Electrical and Computer Engineering
University of California, Davis
Davis, CA, US
zding@ucdavis.edu

***Abstract*—In many packet-switched wireless systems including cellular networks, RObust Header Compression (ROHC) plays an important role in improving payload efficiency by reducing the number of header bits in a link session. However, there are only very few research works addressing the optimized control of ROHC. Our recent studies have demonstrated the advantage of a trans-layer ROHC design that exploits lower layer link status. We have presented a unidirectional ROHC design based on a partially observable Markov decision process formulation that enables the transmitter to decide the header compression level without receiver feedback. The present work considers the physical channel dynamics in an LTE environment and how they affect header decompressor status. Our new model takes into consideration the transport block (TBs) size defined in LTE transmission according to the modulation and coding scheme (MCS). Our novel and practical model can significantly improve the efficiency of the transmission when compared to a traditional timer-based ROHC control.**

***Index Terms*—Packet header, compression, ROHC, Markov decision process, physical channel, cross-layer design.**

I. Introduction

Modern wireless communication systems are increasingly favoring full Internet Protocol (IP) packet-switched architectures [1]. The tremendous growth of wireless data traffics, services, and users continues to push for high spectral efficiency in wireless networks. Traditional focus on bandwidth efficient PHY and MAC layers are no longer sufficient. With ubiquitous IP services, one notices the significant size and the redundancy of packet headers that can impact the overall network bandwidth efficiency. Header compression is a widely adopted technique to reduce the amount of unnecessary packet headers to improve packet payload throughput for the following reasons:

- IP Packet headers are comparable to certain packet payloads [2] in many applications and services, such as interactive games and multimedia streaming.
- Packet headers exhibit high redundancy and are mostly compressible, since many header fields remain unchanged or change predictably during a link session.

RObust Header Compression (ROHC) [3], [4] is a standard responsible for IP header compression in wireless links, which are characterized by high packet error rate and long round trip-time [2]. ROHC has been incorporated in wireless packet-switched cellular networks [5], however, it has thus far only attracted very scant research attention. In particular, existing analysis do not state how to determine compression confidence, with few exceptions as [6] and references therein.

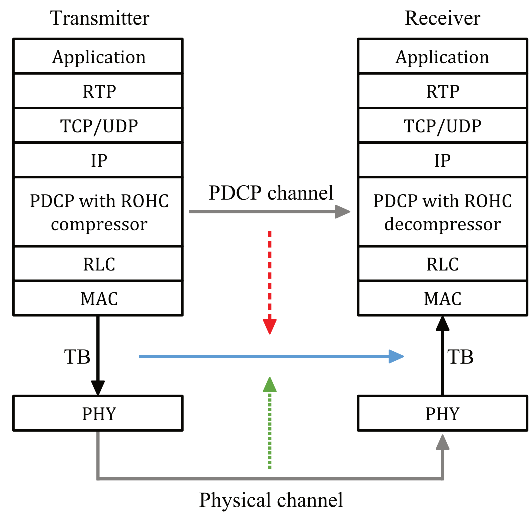

Fig. 1. Protocol stack of a packet switched network. The trans-layer ROHC design proposed in [6] is defined in the PDCP layer, making decisions based on a PDCP-level channel. This work considers the channel model at the TB level, taking into account the physical channel characteristics (green dotted) and the PDCP header compression decisions (red dashed).

Fig. 1 illustrates the protocol stack of a packet switched network. Typically, ROHC resides in the upper Packet Data Convergence Protocol (PDCP) layer and does not use information from other layers. We note that better ROHC decisions could be made if the ROHC compressor can more accurately estimates the success of a PDCP packet and the channel status from lower layer information such as the Channel Quality Indicator (CQI) and HARQ feedback. This trans-layer concept [6] was accompanied by an ROHC design that allows ROHC compressor to exploit information from lower layers in the unidirectional mode (U-mode) of ROHC to achieve higher spectral efficiency.

We note, however, that the trans-layer design suggested in [6] makes header compression decisions at the PDCP level. In other words, the decision of what header compression level to

This material is based on works supported by the National Science Foundation under Grants CNS-1443870 and CNS-1702752. The work of C. Feres is also supported by CONICYT PFCHA/BCH 72170648 scholarship.

978-1-5386-3180-5/18/$31.00 ©2018 IEEE

use at the current time is made for each PDCP packet. This design does not take into consideration the fact that wireless transmission of PDCP packets may require segmentation or aggregation of PDCP packets by lower layers. Thus, the corresponding PDCP header decisions may be better determined by considering the actual channel model that is sensitive to the conditions at the lower layers of the protocol stack, e.g. PHY and MAC layers. Specifically, wireless networks such as LTE do not transmit PDCP packets individually at the PHY layer. Instead, PDCP packets are arranged into transport blocks (TBs). Data in each TB are uniformly subjected to physical channel effects, depending on PHY layer parameters, such as modulation, coding, channel fading, interference levels, and subcarrier channel quality. Clearly, a PDCP-level channel model does not reliably characterize these important attributes aspects and new system models should be considered in order to make better use of lower layer information.

In this work, we present a modified trans-layer U-mode ROHC compressor design that builds on the existing work in [6]. We consider a new TB based header channel model to accurately represent the effect of the channel and other PHY effects. More specifically, we examine the mapping from PDCP packets to TBs in order to propose improvement and adaptations to the PDCP-level channel model formulation. We also develop another mapping from the physical channel to the TB level in order to characterize the error probability of physical channels with different modulation and channel coding schemes (MCS).

The rest of this document is organized as follows. Section II first presents an overview of the ROHC header control in LTE wireless networks. We also introduce the channel model and the necessary assumptions to derive a new POMDP header control. Section III presents the details of the problem formulation and the newly adapted POMDP controller for TB level ROHC. Section IV provides simulation test results. Section V presents conclusions.

II. System Model

A. ROHC Overview

We will focus on the scenario in which a wireless transmitter is operating with a U-mode ROHC compressor, which transmits a PDCP packet stream with compressed headers, whereas a remote receiver with a corresponding ROHC decompressor recovers the compressed packet headers. We shall focus on the compressor design for U-mode ROHC, in which there is no ROHC decompressor feedback to the compressor. This mode is relevant since many wireless services adopt the U-mode ROHC and since ROHC must always start in U-mode before transitioning into other modes (if designed in that way) [3].

Packet headers can be understood as containing a static part that remains unchanged through a link session, and a dynamic part that changes regularly. There are several types of headers in standardized ROHC [3], [4], but as a generalization three types of packet headers can be defined: IR (Initialization and Refresh), which is not compressed at all; First-Order (FO), in which only the static part of the header has been compressed; and Second-Order (SO), in which both static and dynamic parts are compressed. Thus, IR headers are the longest and consume the most bandwidth resources, whereas SO headers are the shortest and the most bandwidth efficient.

The U-mode ROHC compressor decides the level of compression of each packet header, without knowing exactly whether the decompressor is in a state that allows the decompressor to recover the full header from the compressed header based on previously received header information (or ”context” [3, Sec. 2]). Therefore, the compressor always starts with IR packet headers to feed context information to the decompressor. The compressor can decide to apply higher levels of compression (i.e., FO or SO) when it has sufficient confidence that the decompressor has the context to decompress the higher level (FO or SO) header compressions. In short, the compressor should maintain a context synchronization with the decompressor without receiving feedback from the decompressor. Similarly, the compressor also should transition into lower compression levels (IR or FO) when it anticipates the loss of context by the decompressor that may lead to decompression failure. When the wireless receiver fails to receive several data packets at the lower level, for example, then the decompressor is also likely to lose the context to decompress future SO headers, thereby leading to more packet losses in a session. As discussed in [6], the U-mode compressor can either use a timer to periodically send IR or FO packets in anticipation of context loss at the decompressor, or estimate the state of the decompressor so as to make corresponding decisions based on a partially observable Markov Decision Process (POMDP).

With respect to context synchronization, the ROHC decompressor can be modeled as finite-state machine (FSM) with three states:

- Full-Context (FC), in which the decompressor can decode any type of packet;
- Static-Context (SC), in which only the static context is known and one successful FO or IR packet is needed to re-establish full context;
- Non-Context (NC), in which the decompressor requires initialization and can only decode successfully received IR packets, returning to FC.

Each packet failure can be interpreted as a missed context update, which lead to decompressor’s state transitions. A typical header compression algorithm is the Window-based Least Significant Bit (WLSB) [3], which maintains context using a sliding window of length W. Thus, the decompressor can tolerate W consecutive failed packets before needing to re-establish context. Adopting WLSB, the FC state can be expanded into auxiliary states $\mathrm{FC}_0, \mathrm{FC}_1, \ldots, \mathrm{FC}_W$, in which FC_w represents the state in full context after having lost w consecutive packets. The modified FSM of the PDCP ROHC compressor with WLSB encoding is shown in Fig. 2.

B. Dynamic Channel Model

In existing ROHC studies, a Gilbert-Elliot model [7], [8] for the ROHC channel is favored for its simplicity. However, in reality, the wireless channel quality may not be well

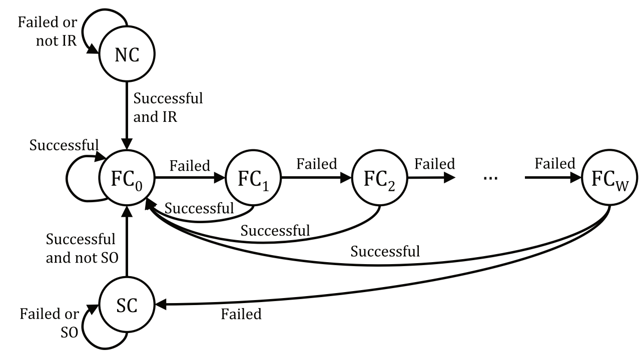

Fig. 2. FSM representation of a PDCP ROHC decompressor with WLSB encoding: the state transition depends on the type of compressed header and whether or not the packet (including header) is successfully received.

captured by only two states. Although [6] applies a general finite state Markov channel (FSMC) to address this issue, the defined FSMC represents a channel at the PDCP level and does not account for the PHY and MAC layer actions in response to a physical channel. For this reason, our approach in this work will stress an FSMC defined at the TB level to properly depict the quality characterization of the wireless channel. Specifically, our wireless channels are mapped into a general finite state Markov chain (FSMC) of K states, where each state denotes a transmission reliability represented by a binary symmetric channel (BSC) with a crossover probability $e_k, k = 1, \ldots, K$. The probability e_k that model state of the channel depends on a number PHY/MAC layer settings, such as link adaptation and HARQ, and the radio channel characteristics, such as signal-to-noise ratio (SNR), co-channel interference, and small or large scale fading [9]. The state transition matrix of this FSMC is denoted by a K-by-K matrix $\mathbf{P}_H$, and the probabilities of successful transmission are denoted by a 1-by-K vector $\boldsymbol{\rho}$, where $\rho_k = 1 - e_k$.

C. Assumptions

Our system model relies on the following practical assumptions:

- A1. IR header, FO head, SO header, and payload in a packet have fixed lengths, denoted by H_0, H_1, H_2 and L_p, respectively, where $H_0 > H_1 > H_2$ reflect different compression levels of IR, FO and SO headers. The total length of IR, FO and SO packets are L_0, L_1, L_2 respectively, with $L_i = H_i + L_P, i = 0, 1, 2$.
- A2. The static part of the headers remains unchanged throughout the lifetime of packet flow, such that the successful transmission of only one IR packet is needed to re-establish static context.
- A3. The transmission delay is relatively negligible.
- A4. The physical channel is a stationary Rayleigh fading channel.
- A5. During a link session, the MCS and the TB size do not change.

Assumptions A1 and A2 are explained in [6], along with all stated considerations therein. Assumption A3 is made for simplicity and to reduce the state space size of this problem, and assumption A4 is commonly used in practice. Assumption A5 can be made because the MCS variation is rather slow compared to channel dynamics, and in high data-rate applications the quality of the transmission link is sufficiently good. This also means that the probability of successful transmission of a TB is content-independent.

III. A TB-level Trans-layer U-mode ROHC Compressor Design

When considering the physical channel and PHY layer settings, three main sources of lower layer information can be identified:

1) The TB size L_{TB} selected by lower layers based on CQI reports and the allotted physical resource blocks (PRBs); The TB size is known at the transmitter based on the MCS and PRBs such that the number of PDCP packets in a TB is also known.
2) Channel quality estimate; This estimate can be computed by analyzing the control signals from PHY/MAC layers. The behavior of the channel estimator can be defined by a matrix $\mathbf{E}_H$, where $E_{H,ij}$ is the probability of getting channel estimate i when j is the true channel.
3) An estimate of last transmitted TB status; This estimate is found from the probabilities of false alarm P_{FA} and missed detection P_{MD}, depending on the known channel quality and transceiver reliability.

The proposed model in [6] has utilized 2) and 3). In this work, we further enable the compressor to take into account the physical channel variation and its effects. Specifically, we develop a more accurate and practical framework to facilitate a new compressor control for ROHC decisions at the TB level. This new development is motivated by two reasons. First, physical channel effect can be reliably represented for each TB in terms of the transport block error rate (BLER) for a given channel. Second, a TB mostly consists of PDCP packets. A TB failure can lead to header losses for all the PDCP packets therein. Thus, it makes sense to consider the role PDCP packet headers play in the TB. The development of the new model will be presented next.

A. Mapping from PDCP Packets into TBs

The conversion of PDCP packets into TBs is not straightforward, as packets can be either segmented or aggregated into different numbers of TBs depending on L_{TB}, which in turn depends on the particular MCS being used for the particular channel quality. Let m be the number of PDCP packets in a TB. We identify two possible scenarios:

1) *Multiple m PDCP packets aggregated in a TB:* In this case, it is obvious that all PDCP packets in a TB are subject to the same TB's transmission conditions. If a TB is received correctly, all PDCP packets within are received correctly; on the other hand, if the TB is lost, all PDCP packets within the TB are lost also. Note that $m = 1$ is merely a special case, which can also be understood as the case studied in [6] while considering

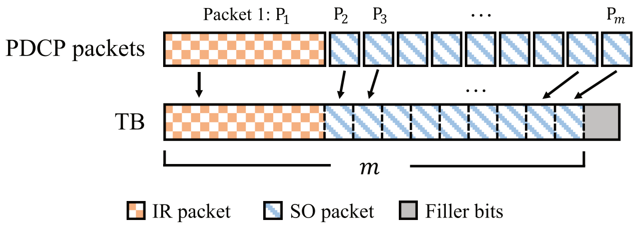

Fig. 3. Construction of a TB from several PDCP packets with header control. The first PDCP packet header may be IR, FO or SO, and all the remaining PDCP packets have SO headers. Filler bits may be used to complete the TB.

a logical PDCP channel. We can safely assume that is an integer $m \geq 1$ with the help of filler bits.

2) *One PDCP packet segmented into multiple TBs* ($m < 1$)*:* In this case, each TB contains a fraction of PDCP packet. Depending on the position of the PDCP packet header, lost TBs may not contain the header and does not necessarily lead to cost context is lost. This case requires a re-definition of what constitutes a lost PDCP packet.

Practically, the case of $m < 1$ is less common in high rate wireless connections and is of less interest since the loss of packet header is highly unlikely. Particularly WLSB ROHC would only lose context when multiple consecutive PDCP packets are lost. When $m < 1$, unless the channel conditions are severely poor, the loss of W consecutive PDCP packets would likely require the loss of W/m consecutive TBs, which would be highly unlikely, particularly given error control mechanisms such as FEC and ARQ. In light of these observations, in the rest of this work we will consider wireless network applications for which $m \geq 1$.

B. Header Control within a TB

With $m \geq 1$, we have already noted that all m PDCP packets in a TB are either successfully received or totally lost under this assumption. With context information, the decompressor will be able to recover all PDCP headers within a successful TB, and may lose context if the TB is lost (depending on the choice of W). Without static context knowledge, the decompressor cannot recover context with a lost TB, but a successful TB containing IR packet allows the decompressor to recover those PDCP packets after the first IR packet while additional IR headers within the TB shall have no effect. A similar argument follows if the decompressor does not have dynamic context knowledge when the TB contains FO packets. Hence, having low order compression headers (IR, FO) in the middle of the TB would be ineffective. For this reason, a more efficient ROHC compressor should only decide to change the ROHC header of the leading PDCP packet of the TB, and automatically assign SO headers to the remaining $m-1$ PDCP packets.

Fig. 3 represents this new control policy, which is a key design feature in our trans-layer ROHC compressor. Under this model, m is selected in such a way that it can fit PDCP packets in one TB with different headers: leading first with an IR, FO or SO packet depending the control decision, followed by $m-1$ SO packets within the TB. Clearly, m should vary depending on the compressor's decision, and is computed as

$$m_i = \left\lceil \frac{L_{\text{TB}} - L_i}{L_2} \right\rceil, \quad i = 0, 1, 2 \tag{1}$$

The corresponding TB-level decompressor can be described by an FSM similar to the one in Fig. 2, with a few adjustments. First, its transitions depend on the first header of the TB and the reception status of the TB. An accurate FSM for this system may have more transition possibilities since a variable number of m_i packets may be lost in each TB. However, for simplicity of analysis and modeling, we shall assume that the decompressor may lose context only when losing an integer number of TBs. Therefore, the window duration W in terms of PDCP packets in Fig. 2 is now approximately replaced with an equivalent TB-level window length W_{TB} defined as

$$W_{\text{TB}} = \left\lfloor \frac{W}{\max_i m_i} \right\rfloor = \left\lfloor \frac{W}{m_2} \right\rfloor. \tag{2}$$

C. FSMC Mapping of Physical Channel

In LTE single antenna transmission mode (SISO), one TB is generated for each Transmission Time Interval (TTI), that has a set duration of 1ms in FDD mode [10]–[12]. Following the methodology presented in [13], the physical channel with Rayleigh fading and an average SNR $\bar{\gamma}$ is mapped into a FSMC with K states of equal average time duration. This approach is preferred over other mapping techniques given the constant time duration of the TBs. The states are defined by the boundaries Γ_k, such that the channel is in state k if $\Gamma_k \leq \gamma \leq \Gamma_{k+1}$, $k = 1, \ldots, K$. Note that we let $\Gamma_1 = 0$ and $\Gamma_{K+1} = \infty$. The resulting transition matrix $\mathbf{P}_H$ represents the channel state transitions for the given TB time duration, the carrier frequency, and the relative velocity of the transceivers.

In this FSMC, the BSC of each state implies that a TB is totally received or totally lost without partial TB recovery, which is consistent with the definition of BSC in Sec. III-A. The crossover probability of each BSC corresponds to the transport block error rate (BLER) in that state. We need to characterize BLER as a function of the SNR, the MCS and the TB size L_{TB} in a AWGN channel:

$$P_{\text{B}}(\gamma, \text{MCS}, L_{\text{TB}}). \tag{3}$$

In the presence of fading with distribution $p(\gamma|\bar{\gamma})$, the channel states have steady-state probabilities $\boldsymbol{\pi}$ given by:

$$\pi_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma|\bar{\gamma}),\, d\gamma \qquad k = 1, \ldots, K \tag{4}$$

And the crossover probability of the BSC of each state can be found as:

$$e_k(\bar{\gamma}, \text{MCS}, L_{\text{TB}}) = \frac{1}{\pi_k} \int_{\Gamma_k}^{\Gamma_{k+1}} P_{\text{B}}(\gamma, \text{MCS}, L_{\text{TB}}) p(\gamma|\bar{\gamma})\, d\gamma \tag{5}$$

The BLER P_{B} can be obtained via tests or simulations for a range of SNR and each MCS, with the knowledge of

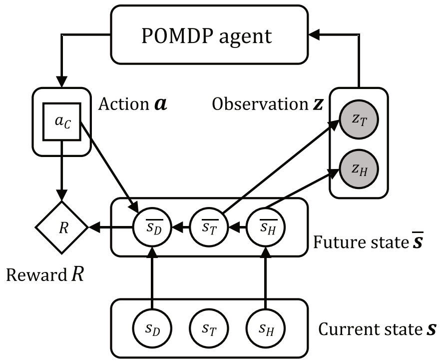

Fig. 4. The DBN representation of the TB-level ROHC system with the trans-layer POMDP formulation. At each timestep, the agent of the POMDP observes the problem via the observable states (in grey), and defines the next action aiming to maximize the long-term reward. In between timesteps, the state of the environment transitions to a new one.

the particular transceiver configuration and parameters. For example, in an LTE system, one could consider the use of turbo coding with specific generator polynomials and interleaver [14], the use of early stopping in turbo decoder [15], CRC validation [16], and HARQ retransmission limit [17]. Once the BLER function is determined, numerical integration can be used to obtain the crossover probability of each state. Curve fitting could also be used to obtain a closed-form expression of the BLER as a function of SNR, MCS, and L_{TB} [18], [19]. In our work, it suffices to use numerical results without having to change simulation setups. Finally, the resulting crossover probabilities are then used to compute $\boldsymbol{\rho}(\bar{\gamma}, \text{MCS}, L_{\text{TB}})$ as explained in Sec. II-B.

D. TB-level POMDP Formulation

With the aforementioned problem forumation at the TB level, we can now define a new U-mode ROHC compressor model for TB. As previously stated, in U-mode the compressor is unaware of the decompressor state. Hence, our design should allow the compressor to estimate the decompressor state based on information from lower layers at the transmitter.

We formulate our header control as a partially observable Markov decision process (POMDP), in which an agent updates its belief on the state of the dynamic system from partial information, and makes a decision regarding which optimum action to take. The POMDP formulation is shown as a dynamic Bayesian network (DBN) in Fig. 4, and is defined by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R \rangle$. We explain the POMDP tuple below.

- $\mathcal{S}$ is the set of states $\mathbf{s} = (s_D, s_T, s_H) \in \mathcal{S} = \mathcal{D} \times \mathcal{T} \times \mathcal{H}$, where $s_D \in \mathcal{D} = \{0, \dots, W_{\text{TB}} + 2\}$ is the state of the decompressor corresponding to $\text{FC}_0, \text{FC}_1, \dots, \text{FC}_{W_{\text{TB}}}$, SC and NC respectively; $s_T \in \mathcal{T} = \{0, 1\}$ is the reception status, where 0 means failure and 1 means successful; and $s_H \in \mathcal{H} = \{1, \dots, K\}$ is the channel state in the transmission.
- $\mathcal{A}$ is the set of actions, that contains the type of ROHC header $a_C \in \mathcal{A} = \{0, 1, 2\}$ (corresponding to IR, FO and SO, respectively) decision for the first packet in the TB.
- $\mathcal{Z}$ is the set of observations $\mathbf{z} = (z_T, z_H) \in \mathcal{Z} = \mathcal{T} \times \mathcal{H}$, where $z_T \in \mathcal{T}$ is the reception status observed by the compressor, and $z_H \in \mathcal{H}$ is the observed channel state.
- $T(\mathbf{s}, a, \bar{\mathbf{s}}) = p(\bar{\mathbf{s}}|\mathbf{s}, a)$ is the transition probability from $\mathbf{s}$ to $\bar{\mathbf{s}}$ given action a_C, and is defined as:
$$T(\mathbf{s}, a_C, \bar{\mathbf{s}}) = p(\bar{s}_H|s_H)p(\bar{s}_T|\bar{s}_H)p(\bar{s}_D|s_D, a_C, \bar{s}_T) \quad (6)$$
in which the conditional probabilities include
  - $p(\bar{s}_H|s_H)$ as the channel state transition, given by the Markov matrix $\mathbf{P}_H$.
  - $p(\bar{s}_T|\bar{s}_H)$ as the reception status distribution, obtained as $\boldsymbol{\rho}$.
  - $p(\bar{s}_D|s_D, a_C, \bar{s}_T)$ as the decompressor transition probability, defined by the FSM model of the TB-level decompressor described in Sec III-A.
- The observation function $O(\bar{\mathbf{s}}, a_C, \mathbf{z}) = p(\mathbf{z}|\bar{\mathbf{s}}, a_C)$ is the probability of observing $\mathbf{z} \in \mathcal{Z}$ in state $\bar{\mathbf{s}}$ after executing action a_C, and is defined as:
$$O(\mathbf{s}, a_C, \bar{\mathbf{s}}) = p(z_H|\bar{s}_H)p(z_T|\bar{s}_T) \quad (7)$$
where:
  - $p(z_H|\bar{s}_H)$ represents the channel estimation function, given by $\mathbf{E}_H$.
  - $p(z_T|\bar{s}_T)$ represents the reception status estimation function, which depends on P_{FA} and P_{MD}.
- The reward $R(\mathbf{s}, a, \bar{\mathbf{s}})$ yields the instantaneous reward obtained by moving from $\mathbf{s}$ to $\bar{\mathbf{s}}$ given action a. Note that, assuming $m \geq 1$ PDCP packets in a TB and noting that the compressor only decides on the first PDCP header of the TB, the transmission efficiency η can be defined as:
$$\eta = \mathbb{E}\left\{ \frac{\sum_{i=0}^{\infty} m_{a_C[i]} \cdot L_p \cdot \mathbf{1}[s_D[i] = 0]}{\sum_{i=0}^{\infty} L_{\text{TB}}} \right\} \quad (8)$$
where $\mathbf{1}[expr]$ is the indicator function, that takes value 1 if $expr$ is true and 0 otherwise. By approximating η with the expected discounted sum of instantaneous transmission efficiency,
$$\tilde{\eta} = \sum_{i=0}^{\infty} \gamma^i \mathbb{E}\left\{ \frac{m_{a_C[i]} \cdot L_p \cdot \mathbf{1}[s_D[i] = 0]}{L_{\text{TB}}} \right\}, \quad (9)$$
then the instantaneous reward function for a classical POMDP formulation can be defined as
$$R(\mathbf{s}, a, \bar{\mathbf{s}}) = \frac{m_{a_C[t]} \cdot L_p}{L_{\text{TB}}} \cdot \mathbf{1}[\bar{s}_D[t] = 0]. \quad (10)$$

The solution of the POMDP is a policy that maps the agent's belief of the state into actions, with the goal of maximizing the long-term reward. POMDP problems are generally complex to solve exactly. However, efficient POMDP solvers are available [20]–[23]. We adopt the SARSOP algorithm [20] to solve the TB-level POMPD U-mode ROHC compressor formulation, which can provide solutions over a reasonable amount of time even using a general purpose PC. Note that the POMDP needs to be solved once for a given set of system settings. Thus, the formulation can be solved during the ROHC negotiation process or even offline (using policy look-up) for feasibility in real-world applications.

IV. Simulation Results

In this section, we present simulation results to illustrate the performance advantage of the proposed POMDP ROHC header control in comparison with an optimized U-mode ROHC compressor utilizing timer as presented [6]. We also apply aggregation of PDCP packets within a TB.

A. Test Setup

Unless stated otherwise, the ROHC system is modeled with the settings listed in Table I. The channel parameters have been selected for a common single-antenna channel in LTE. The number of states $K = 16$ is selected such that the constant c_k in [13] is within an adequate range that ensures transitions only to adjacent states as well as equal BLER for every TB transmitted in the state. The header and payload lengths are described in [6, Sec. VI-A], whereas the WLSB parameter W follows what was used in [24].

We let d_0, d_1 and d_2 denote the mean durations of IR, FO and SO packets, respectively [6, Figure 4]. Traditional ROHC control in U-mode follows a timer based approach. To make fair comparisons with our TB-level compressor, we need to adjust the optimized timer-based compressor of [6] for a TB-level channel to affect m PDCP packets. A typical timer-based compressor which transmits one IR packet and N segments of FO and SO packets alternatingly within a period. Defining $d_0 = 1$, we follow the optimization of mean durations d_1 and d_2 [6] by selecting N as a design parameter, which can be seen as a static context update timer. The alternative timer-based compressor shall select the header of each PDCP packet, and applies another independent procedure to aggregate m packets into a TB of size L_{TB} bits without knowledge of their contents or header types. This means that an IR packet may be any one of the m packets within TB and is not necessarily in the beginning of the TB. If the decompressor receives such a TB when the decompressor is in the NC state, then it shall be unable to recover context until it finds the IR packet. In this case, the decompressors shall lose all packets in the TB that precedes the IR packet. A similar argument applies for FO packets within the TB when the decompressor is in SC state. This alternative, timer-based ROHC header control is consistent with the basic principle of timer-based ROHC header control. Its key difference with our proposed ROHC compressor lies in the fact that the proposed POMDP decision only needs to select the first PDCP header of the TB, thereby improving the payload transmission efficiency.

With respect to the computation complexity, all POMDP instances (with different parameters, such as average SNR) are solved on a basic PC (Intel Core-i7 4790 CPU and 16GB DDR3 memory) within 30 seconds. The results show that the maximum gap between value function bounds reported by the SARSOP algorithm is 1.42%.

B. Performance Results

In Fig. 5, we demonstrate the performance gain of our proposed TB-level POMDP compressor over the timer-based compressor with respect to the average SNR of the channel.

TABLE I
DEFAULT SIMULATION SETTINGS FOR OUR ROHC DESIGN TESTS.

Channel Model	Rayleigh fading, Single antenna $f_c = 1.9\,\text{GHz}$, $v = 5\,\text{km/h}$, $K = 16$
Header/Payload lengths	$H_0 = 59$, $H_1 = 15$, $H_2 = 1$, $L_p = 20$ bytes
WLSB	$W = 5$ ($W_{\text{TB}} = 0$)
Timer Compressor	Optimized d_1 and d_2, no slow-start $d_0 = 1$, $N = 5$
POMDP Compressor	$P_{\text{FA}} = P_{\text{MD}} = 0.1$, $\mathbf{E}_H = I_K$
TB size	$L_{\text{TB}} = 5736$ ($I_{\text{TBS}} = 6$, $N_{\text{PRB}} = 55$ [12, Sec. 7.1.7])

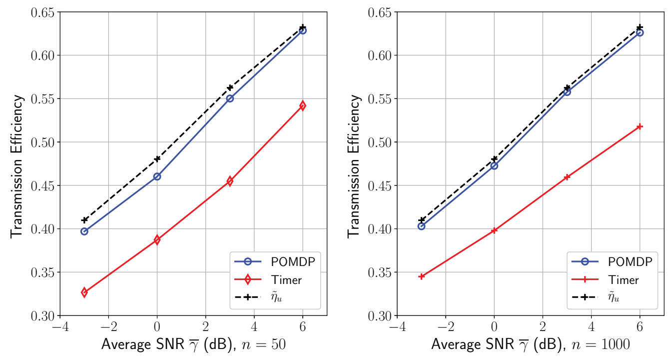

Fig. 5. The empirical efficiency η versus the average SNR of the channel, both transient and steady-state.

We show two different timesteps of $n = 50$ and $n = 1000$, respectively, to illustrate the transient and steady-state efficiency gains. We also provide the asymptotic expected efficiency of the POMDP compressor as a reference, defined as the efficiency when selecting SO headers as the first PDCP header in a TB assuming no decompression failure,

$$\tilde{\eta}_u = \boldsymbol{\rho}^T \boldsymbol{\pi} \frac{m_2 L_p}{L_{\text{TB}}}. \tag{11}$$

In both transient and steady state situations, the POMDP compressor achieves clear performance gain over the timer-based compressor by as much as 10%. Moreover, the POMDP compressor efficiency is close to η_u, considering decompression failures and TBs with longer initial packet headers.

Note that $\tilde{\eta}_u$ is not necessarily applicable to the timer-based compressor, because different selections of N also change the number of PDCP packets in a TB m_{timer}, which, in turn, changes the expected efficiency of this compressor. The results in Fig. 6 illustrate the efficiency of the timer-based compressor with respect to different values of N. The results are shown in comparison with the efficiency of the newly proposed POMDP compressor. It can be seen that the new compressor is superior for every selection of N. In fact, larger N does not always lead to performance gain since there are fewer IR packets.

Finally, we present the efficiency of our compressor when subject to estimation errors in Fig. 7. We define $P_{e,T} = P_{\text{FA}} = P_{\text{MD}}$ as the probability of erroneously estimating the transmission status. The probability of error in channel state estimation $P_{e,H}$ is defined by obtaining an adjacent state

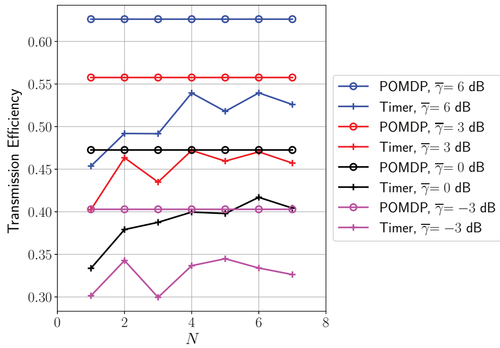

Fig. 6. The empirical efficiency η versus the static context update timer N.

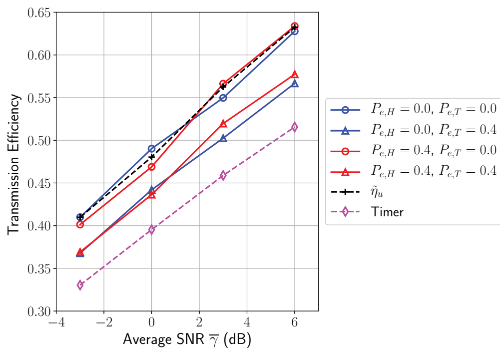

Fig. 7. The empirical efficiency η of the POMDP compressor versus average SNR, for different values of the channel estimation error probability α. The timer-based compressor is shown for comparison.

as estimate (equiprobably selected if there is more than one adjacent state, for simplicity), that is,

$$\begin{aligned} P_{e,H} &= P(z_2|s_1) = P(z_{K-1}|s_K) \\ &= 2P(z_{k-1}|s_k) = 2P(z_{k+1}|s_k), \quad 1 \leq k \leq K \end{aligned} \tag{12}$$

The efficiency does not deviate significantly from η_u for different values of $P_{e,H}$, which means that our compressor is robust against channel estimation errors. The POMDP compressor is more sensitive to transmission status estimation errors, although it still performs better than the timer-based compressor (with settings given by Table I).

V. Conclusion

This work presents a trans-layer design framework for a U-mode ROHC compressor that considers the practical transport layer and makes optimized decisions at the transport block level. This framework can directly incorporate the physical channel conditions in the 4G-LTE cellular networks. Our design formulates the ROHC header control optimization into a partially observable Markov decision process (POMDP). We map the physical channel and the PDCP packets into the transport block level POMDP framework. Our proposed compressor can operate without complete knowledge of the decompressor state in the U-mode and without relying on immediate wireless channel quality report. When compared with the traditional timer-based ROHC controller without taking advantage of lower layer information as we do, the proposed POMDP controller significantly improves the packet delivery efficiency for a wide range of design parameters and channel SNR.

References

[1] Federal Communications Commission (FCC), *Connecting Ammerica: The National Broadband Plan*, 2010. [Online]. Available: https://www.fcc.gov/general/national-broadband-plan

[2] “The concept of robust header compression, ROHC,” *Effnet AB whitepaper*, 2004.

[3] C. Bormann, C. Burmeister, M. Degermark, H. Fukushima, H. Hannu, L.-E. Jonsson, R. Hakenberg, T. Koren, K. Le, Z. Liu, A. Martensson, A. Miyazaki, K. Svanbro, T. Wiebke, T. Yoshimura, and H. Zheng, “RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed,” RFC 3095, Jul. 2001.

[4] K. Sandlund, G. Pelletier, and L.-E. Jonsson, *The RObust Header Compression (ROHC) framework*, RFC 5795, Mar. 2010.

[5] *Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification*, 3GPP Tech. Specification TS 36.323 (v13.2.1), 2016.

[6] W. Wu and Z. Ding, “On Efficient Packet-Switched Wireless Networking: A Markovian Approach to Trans-Layer Design and Optimization of ROHC,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4232–4245, July 2017.

[7] L. Badia, N. Baldo, M. Levorato, and M. Zorzi, “A Markov framework for error control techniques based on selective retransmission in video transmission over wireless channels,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 488–500, Apr. 2010.

[8] E. O. Elliott, “Estimates of error rates for codes on burst-noise channels,” *The Bell System Technical Journal*, vol. 42, no. 5, pp. 1977–1997, Sep. 1963.

[9] C. Sahin, L. Liu, and E. Perrins, “Coding Across Finite Transport Blocks in Modern Wireless Communication Systems,” *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4184–4197, Dec 2014.

[10] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, 3GPP Tech. Specification TS 36.211 (v13.0.0), 2016.

[11] *Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding*, 3GPP Tech. Specification TS 36.212 (v13.2.0), 2016.

[12] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures*, 3GPP Tech.l Specification TS 36.213 (v13.2.0), 2016.

[13] Q. Zhang and S. A. Kassam, “Finite-state Markov model for Rayleigh fading channels,” *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1688–1692, Nov 1999.

[14] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution*. Wiley Online Library, 2015.

[15] J. F. Cheng and H. Koorapaty, “Error Detection Reliability of LTE CRC Coding,” in *IEEE 68th Vehicular Technology Conf.*, Sept 2008, pp. 1–5.

[16] M. El-Khamy, J. Lee, and I. Kang, “Detection Analysis of CRC-Assisted Decoding,” *IEEE Commun. Lett.*, vol. 19, no. 3, pp. 483–486, Mar. 2015.

[17] A. Masaracchia, R. Bruno, A. Passarella, and S. Mangione, “Analysis of MAC-level throughput in LTE systems with link rate adaptation and HARQ protocols,” in *IEEE 16th Int. Symp. World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, June 2015, pp. 1–9.

[18] S. Lembo, “Modeling BLER performance of punctured turbo codes,” Master’s thesis, Aalto University, Espoo, Finland, May 2009.

[19] M. M. Anis, “Performance Prediction of a Turbo-coded Link in Fading Channels,” Master’s thesis, Aalto University, Espoo, Finland, Mar. 2010.

[20] APPL: Approximate POMDP planning toolkit. [Online]. Available: http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/

[21] A. R. Cassandra. pomdp-solve v5.4. [Online]. Available: http://www.pomdp.org/code/

[22] E. Patrick. POMDPy. [Online]. Available: http://pemami4911.github.io/POMDPy/

[23] H. Kurniawati, D. Klimenko, J. M. Song, K. Seiler, and V. Yadav. TAPIR: Toolkit for approximating and Adapting POMDP solutions In Real time. [Online]. Available: http://robotics.itee.uq.edu.au/ hannakur/dokuwiki/doku.php?id=wiki:tapir

[24] S. Kalyanasundaram, V. Ramachandran, and L. M. Collins, “Performance analysis and optimization of the window-based least significant bits encoding technique of ROHC,” in *2007 IEEE Global Telecommun. Conf.*, Nov. 2007, pp. 4681–4686.