

Low Complexity Header Compression with Lower-Layer Awareness for Wireless Networks

Carlos Feres, *Student Member, IEEE*, Zhi Ding, *Fellow, IEEE*
Department of Electrical and Computer Engineering
University of California, Davis
Davis, CA, US

Abstract—Packet-switched wireless networks such as 4G and 5G cellular systems apply RObust Header Compression (ROHC) to reduce PDCP header length and improve payload efficiency. Our recent works have demonstrated the benefit of applying a trans-layer approach that exploits lower layer information in ROHC control based on a partially observable Markov decision process (POMDP) formulation. The benefit of the POMDP solution comes at significant computation complexity beyond existing ROHC. The present work focuses on simplicity by designing ROHC compressors with lower layer awareness including channel adaptive transport block size as a result of link adaptation present in many wireless networks. Our new models directly address practical implementation and can deliver transmission efficiency close to an optimized POMDP compressor.

Index Terms—Packet header, header compression, ROHC, Markov decision process, link adaptation.

I. INTRODUCTION

Modern wireless communication networks are converging toward full Internet Protocol (IP) packet-switched architectures [1]. Considering the growing number of users and services, traditional approaches at PHY and MAC layers for improving bandwidth efficiency are already near the capacity limit. A more comprehensive approach that transcends all network layers is desired. In particular, IP is characterized by long and redundant packet headers can sacrifice bandwidth efficiency. As a result, header compression schemes are widely adopted to reduce headers to improve payload efficiency [2]. In particular, RObust Header Compression (ROHC) [3], [4] is a standardized technology widely adopted in wireless packet-switched networks such as 4G-LTE [5] and 5G-NR [6]. However, ROHC has only received scant research attention, most of which has been focused on improvements or analysis under very specific conditions [7]–[10]. In contrast, our previous works [11]–[14] have presented novel Markovian design frameworks for optimizing ROHC systems in terms of bandwidth efficiency, with a general approach and emphasis in lower-layer information.

In this work, we present new trans-layer U-mode ROHC compressor designs with different levels of complexity that generalize our previous proposal [13]. Specifically we consider a new channel model that represents the effect of physical channel dynamics, link adaptation, and other PHY/MAC layer effects at the transport block (TB) level, along with the necessary considerations to map the PDCP packet level decisions

of ROHC compressors to the TB-level. We also examine new formulations that yield simpler ROHC designs, while suffering little performance loss versus our optimized POMDP compressor.

The rest of this document is organized as follows. Section II presents an overview of the ROHC header control in 4G and 5G wireless networks, and introduces the channel model and the necessary assumptions in our new ROHC control framework. Section III presents the ROHC compressor formulation along with the features and details that facilitate low complexity ROHC decision. Section IV provides simulation test results to demonstrate the performance advantage of our solutions, before the conclusions of Section V.

II. SYSTEM MODEL

A. ROHC Overview

This work focuses on the scenario in which a wireless transmitter operates a U-mode ROHC compressor without a feedback from decompressor. The U-mode ROHC compressor transmits a PDCP packet stream of compressed and uncompressed headers and the corresponding decompressor at the receiver recovers the compressed headers. In particular, we will study controlling the U-mode ROHC compressor without direct feedback on decompressor's state. Many low latency wireless services adopt the U-mode ROHC. Furthermore, any ROHC system must always start in U-mode before transitioning into other feedback-based modes [3], [4].

IP packet headers contain a static field that remains unchanged during a link session, and a dynamic field that changes regularly with packets. Although there are several types of headers in standardized ROHC with different levels of compression [3], [15], we consider the simple and common case with three types of packet headers: IR (Initialization and Refresh), which is uncompressed; First-Order (FO), which has a compressed static part and uncompressed dynamic part; and Second-Order (SO), in which both static and dynamic parts are compressed. Therefore, IR headers are the longest and require the most bandwidth, whereas SO headers are the shortest use the least bandwidth.

Both ROHC compressor and decompressor aim to maintain a shared header information, or "context", throughout the link session. A U-mode ROHC compressor has to decide the compression level of each packet header without feedback, i.e., without knowledge about the decompressor state and whether it is able to decode a compressed header or not. Hence, the U-mode compressor must start using IR packet headers 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.

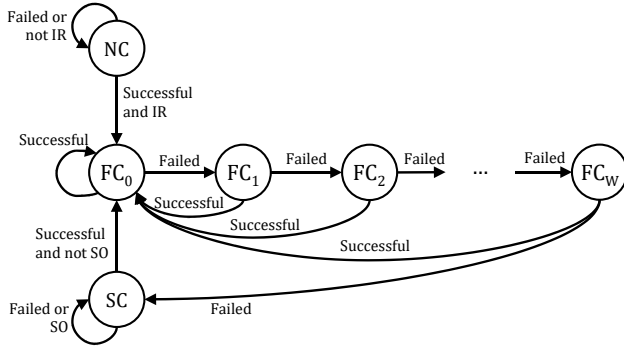


Fig. 1. FSM representation of a PDCP ROHC decompressor with WLSB encoding: the state transition depends on the type of compressed header and whether or not the packet (including header) is successfully received.

establish context at the decompressor. The compressor can then decide to apply higher levels of compression (i.e., FO or SO) once it is confident of the decompressor's ability to decompress SO or FO headers with the provided context. Correspondingly, the compressor can take 3 actions and the space of its last action contains 3 states: IR, FO, and SO according to the selected compression level. At a given time, the decompressor is in one of 3 states: NC (no context), RC (repair context), and FC (full context). Depending on its state, the ROHC decompressor can only receive IR header in NC, either FO or IR header in RC, and any of the 3 headers in FC.

Traditional U-mode compressor is based on a timer mechanism that periodically update the context using IR, in anticipation of context loss [3]. Header decompression is usually based on the Windowed Least-Significant Bit (WLSB) algorithm [3], which maintains context within W consecutive packet losses. A ROHC decompressor with WLSB can be modeled as a finite-state machine with $W + 2$ states as the one depicted in Fig. 1 [13].

B. Dynamic CQI-based Channel Model

Even though most ROHC-related works adopt a Gilbert-Elliott channel model, true wireless channels and the PHY/MAC layers exhibit more dynamic characteristics that may not be captured well by a two-state model. Among others, an important feature of wireless networks is link adaptation which adjust the modulation and coding scheme (MCS) throughout the link session to keep its packet error rate below a certain level. Previously, we have neglected the link adaptation effect by assuming its slow variation when compared with the rate of channel variations. To model the channel variation and link adaptation more accurately, we define a new finite-state Markov channel (FSMC) model to address the varying MCS transmission during link session.

Consider a finite state channel model with K states that corresponds to K the possible CQI values of the wireless link. Each channel state represents a bit-level symmetric channel (BSC) with a crossover error probability $e_k, k = 0, \dots, K - 1$. Moreover, each channel state will be defined by the parameters selected for that particular CQI value, i.e., the transmitter in

state k uses the modulation and coding settings given by CQI value k , which in turn provides the TB size B_k (in bytes). Additionally, the use of CQI reports as channel states rely exclusively on the existing CQI-reporting mechanism, and thus avoids the requirement of modeling particular channel characteristics and transmission conditions and the implementation of corresponding estimations where needed.

The error probability e_k for each state can be found through test or simulation. For LTE and 5G NR, the target BLER of 10% for each MCS can be used as a worst-case scenario. Let state $k = 0$ denote a no-transmission state with $B_0 = 0$. This FSMC is characterized by a $K \times K$ state transition matrix $\mathbf{P}_H$ and the probabilities of successful transmission are denoted by a $1 \times K$ vector $\boldsymbol{\rho}$, where $\rho_k = 1 - e_k$. Under certain channel conditions, not all CQI states are possible or reachable, leading to zero rows in $\mathbf{P}_H$.

C. Assumptions

Our system model relies on the following practical assumptions:

- A1. IR header, FO header, SO header, and payload in a packet have fixed lengths, denoted by H_0, H_1, H_2 and L_p , respectively, where $H_0 > H_1 > H_2$ correspond to compression levels of IR, FO and SO headers, respectively. The total length of IR, FO and SO packets are L_0, L_1, L_2 respectively, with $L_m = H_m + L_p, m = 0, 1, 2$.
- A2. The static part of source headers remains unchanged throughout the lifetime of packet flow, such that successful transmission of only one IR packet is needed to re-establish static context.
- A3. The dynamic compressibility of source headers follows a Markov Decision Process, such that some headers are not compressible. In other words, apart from context damage due to packet loss, an IR or FO header needs to be transmitted to re-establish context if a packet source header is not compressible.
- A4. The transmission delay is negligibly short.

In our previous work [13], all the stated assumptions were also provided, along with corresponding practical considerations. In the current formulation, we no longer assume constant MCS and TB size in consideration of link adaptation in wireless links.

III. TRANS-LAYER U-MODE ROHC COMPRESSION

A. Initial considerations

As specified in both LTE and NR standards [6], [16], [17], IP packets undergo ROHC compression by the PDCP before entering the RLC layer, which reorders and adds RLC headers to form RLC PDUs that are used by the MAC layer to form a TB. Although there are some differences among different standards, the general process is essentially the same from the perspective of ROHC design: PDCP packets are concatenated or segmented in the process such that a large TB can be composed of multiple concatenated PDCP packets, or a PDCP

packet can be segmented into several small TBs, depending on the MCS used in link adaptation.

For our investigation and in general, the low data rate scenario where a single packet is segmented in multiple TBs is of less interest, as a WLSB-based ROHC requires several consecutive TB failures to lose context synchronization. We focus instead on high data rate links in which one TB contains multiple PDCP packets. Additionally, we will neglect the segmentation of one packet in two successive TBs by assuming that multiple full PDCP packets reside in each TB, in addition to some filler bits, as depicted in Fig. 2.

In our previous work [13], we proposed a header compression controller that decides the compression level of only the first header for each TB whereas the subsequent packets in the TB use fully compressed SO headers. This header control policy offers increased efficiency as more packets and higher payload will be carried by the TB as a result. This control scheme is based in the fact that a TB is either successfully received in its entirety or is totally lost. Thus, the decompressor operates as follows:

- In NC state, it establishes context only with a successfully received TB that starts with the IR header and can decode subsequent packet headers. It will be unable to do so if the TB is either lost or contains no IR headers;
- In SC state, it can recover the dynamic context with a successfully received TB that starts with an IR or FO header and can decode subsequent packet headers. It may lose static context if the TB is either lost contains only SO headers;
- In FC state, it can recover all PDCP headers in a successful TB, and otherwise may lose context depending on number of PDCP packets in the TB and the value W .

With a CQI-based FSMC that considers link adaptation, assuming multiple PDCP packets per in a TB that uses header control as depicted in Fig. 2, the number of PDCP packets per TB is denoted by a $K \times 3$ constant matrix $\mathbf{M}$ in which each element $m_{k,a}$ represents the integer number of PDCP packets per TB for CQI index k (given by the channel state), and the first packet header in the TB a (with all other packets having SO headers), given by

$$m_{k,a} = \left\lceil \frac{B_k - L_a}{L_2} \right\rceil. \quad (1)$$

To conclude, we can obtain a rather simple TB-level approximation of a traditional PDCP-level ROHC decompressor with WLSB. If TBs are either successfully received or totally lost, it makes sense to assume that the decompressor loses context after an integer number of lost TBs. Thus, we define an approximated TB-level window W_{TB} which depends on the number of packets per TB, although to avoid a varying window parameter in our formulations, we use the approximation:

$$W_{\text{TB}} = \left\lfloor \frac{W}{\min_{k \neq 0} \max_a m_{k,a}} \right\rfloor. \quad (2)$$

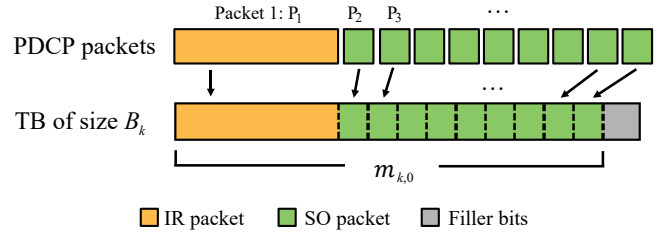


Fig. 2. Construction of a TB of variable size from several PDCP packets with header control. The first PDCP packet header may be IR, FO or SO, and all the remaining PDCP packets have SO headers (with IR-leading packet depicted in this particular case). Filler bits may be used to complete the TB.

B. POMDP Compressor

We reformulate the POMDP compressor discussed previously [13] into a new POMDP formulation by considering link adaptation. The new POMDP is defined by the tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R \rangle$, in which each element of the tuple has the same function as in [13, Sec. III-D], with the exception that

- For the set of states $\mathcal{S}$, the channel state $s_H \in \mathcal{H} = \{0, \dots, K-1\}$ is shifted to directly map the CQI report as the channel state index.
- In this case, the efficiency of the transmission (defined as successfully decompressed payloads over transmitted TBs) now considers the variable TB size:

$$\eta = \mathbb{E} \left\{ \frac{\sum_{t=0}^{\infty} m_{s_H[t], a_C[t]} \cdot L_p \cdot \mathbf{1}[s_D[t] = 0]}{\sum_{t=0}^{\infty} B_{s_H[t]}} \right\}, \quad (3)$$

in which $\mathbf{1}[expr]$ is an indicator function, that takes value 1 if $expr$ is true and 0 otherwise. Here for notation simplicity we let $0/0 = 0$. We approximate η with the expected discounted sum of instantaneous transmission efficiency:

$$\tilde{\eta} = \sum_{t=0}^{\infty} \gamma^t \mathbb{E} \left\{ \frac{m_{s_H[t], a_C[t]} \cdot L_p \cdot \mathbf{1}[s_D[t] = 0]}{B_{s_H[t]}} \right\}, \quad (4)$$

Then the new instantaneous reward function $R(s, a, \bar{s})$, obtained by moving from s to $\bar{s}$ given action a , is

$$R(s, a, \bar{s}) = \frac{m_{\bar{s}_H[t], a_C[t]} \cdot L_p}{B_{\bar{s}_H[t]}} \cdot \mathbf{1}[\bar{s}_D[t] = 0]. \quad (5)$$

The solution of the POMDP is a decision policy $\mathbf{p}$ that maps the compressor's belief of the state $\mathbf{q}(s), s \in \mathcal{S}$ into actions, i.e. $a = \mathbf{p}(\mathbf{q})$, aiming to maximize the reward in the long-term. In general, POMDP problems are difficult to solve exactly. However, numerically efficient POMDP solvers are available [18]–[21].

C. TB-level Timer-based Compressor

Although the POMDP policy should yield close-to-optimal results and only needs to be solved offline once for given system settings for future table lookup, in practice it requires knowledge of the channel dynamics to construct an adaptive model. This could be quite impractical given the wide range

of transmission conditions and the prodigious memory size of such a look-up policy. Thus, it may be desirable to construct a timer-based ROHC compressor that is aware of lower-layer information and dynamics. Such a deterministic and periodic policy would be easier to implement and shall be known as TB-level Timer-based (TB-timer) compressor.

However, it is difficult to analytically obtain a direct relationship between the efficiency of the compressor and the timer parameters. Instead, one may have to rely on empirical results from Monte-Carlo simulations. On the other hand, an exact Markov model of the compressor-channel-decompressor system may have a significantly large state space if the timer period is quite large. Therefore, by adopting the header control for the TB structure described in Section III-A, we will approximate the ROHC compressor as a Markov process with 3 states representing the first header decision of a TB to optimize the Markov compressor. We shall then map the optimized results to a timer-based design.

To begin, this Markov chain requires a 3×3 state transition matrix $\mathbf{P}_C$, whose entry $P_{C,ij}$ is the transition probability from state i to state j for $i, j = 0, 1, 2$, each respectively representing the selection of IR, FO or SO header. Each column of $\mathbf{P}_C$ is completely defined by 2 parameters. Thus, the steady-state behavior of any U-mode compressor can be approximated by 6 parameters: the fraction and average duration of IR-, FO-, and SO-leading TBs, denoted f_{s_C} and d_{s_C} respectively, with $s_C = 0, 1, 2$ denoting IR, FO and SO. Because $\sum_{s_C=0}^2 f_{s_C} = 1$, the U-mode compressor's degree of freedom is 5 [11]. To derive a one-to-one mapping between timer and Markovian parameters, we realize that transitions from IR to FO headers in TBs do not assist in context synchronization except for some unnecessary overhead. Thus, we shall let $P_{C,10} = 0$ such that $\mathbf{P}_C$ has 5 parameters. The fraction vector $\mathbf{f} = [f_0, f_1, f_2]$ corresponds to the steady-state distribution of $\mathbf{P}_C$, i.e.,

$$\mathbf{f} = \mathbf{f} \mathbf{P}_C. \quad (6)$$

The average duration d_i of consecutive type- i leading TBs can be defined as

$$d_0 = \frac{1}{P_{C,02}} \quad d_1 = \frac{1}{P_{C,10} + P_{C,12}} \quad d_2 = \frac{1}{P_{C,20} + P_{C,21}}. \quad (7)$$

The mapping has a corresponding feasibility region of

$$\mathcal{F} = \{(d_0, d_1, d_2, f_0, f_1, f_2) : d_0 \geq 1, d_1 \geq 1, d_2 \geq 1, \frac{f_2}{d_2} \geq \frac{f_1}{d_1}, \frac{f_2}{d_2} \geq \frac{f_0}{d_0}, \frac{f_0}{d_0} + \frac{f_1}{d_1} \geq \frac{f_2}{d_2}\}. \quad (8)$$

Thus, we can optimize the parameters of the Markov compressor before mapping the obtained optimal set onto the timer parameters. In particular, the deterministic timer-based structure shown in Fig. 3 fits the above description and satisfies the feasibility constraints.

The system has a state space $\mathcal{S} = \mathcal{S}_C \times \mathcal{S}_H \times \mathcal{S}_D$ representing the product space consisting of compressor state space $\mathcal{S}_C$, channel state space $\mathcal{S}_H$ and decompressor state space $\mathcal{S}_D$.

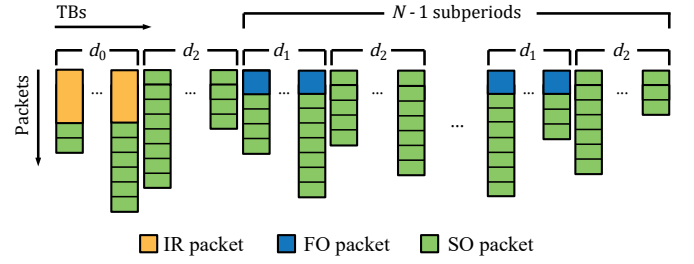


Fig. 3. TB-level timer-based structure with header control. This periodic structure consists of N subperiods, the first having d_0 IR-leading TBs followed by d_2 SO-leading TBs, and then $N-1$ subperiods of d_1 FO-leading TBs followed by d_2 SO-leading TBs. Additionally, the TBs may have different sizes depending on the CQI report.

The full system state is simply $\mathbf{s} = (s_C, s_H, s_D) \in \mathcal{S}$. The transmission status s_T depends on the compressor state s_C and the new channel state s'_H . The transition probability of the system is given by

$$p(\mathbf{s}'|\mathbf{s}) = p(s'_C|s_C)p(s'_H|s_H)p(s'_D|s_D, s_C, s'_H) \quad (9)$$

where the compressor and channel transition probabilities $p(s'_C|s_C)$ and $p(s'_H|s_H)$ are defined by $\mathbf{P}_C$ and $\mathbf{P}_H$, and

$$p(s'_D|s_D, s_C, s'_H) = p(s'_D|s_D, s_T = 0, s_C)p(s_T = 0|s'_H) + p(s'_D|s_D, s_T = 1, s_C)p(s_T = 1|s'_H) \quad (10)$$

where $p(s_T = 1|s'_H) = \rho_{s'_H}$ and $p(s'_D|s_D, s_T = i, s_C) \in \{0, 1\}$ is defined by the FSM representation of the decompressor in Fig. 1.

The asymptotic efficiency of the system can be described as the average number of successfully decompressed packet payloads (which can be derived using total probability law with the steady-state probabilities of compressor, channel and decompressor in each state) over the average TB size. Thus, we obtain the optimization problem

$$\begin{aligned} \max \quad & \eta = \frac{\pi_{D,0} L_p \sum_{s_H=0}^{K-1} \sum_{s_C=0}^2 \pi_{H,s_H} \pi_{C,s_C} m_{s_H,s_C}}{\sum_{s_H=0}^{K-1} \pi_{H,s_H} B_{s_H}} \\ \text{s.t.} \quad & P_{C,02}, P_{C,10}, P_{C,12}, P_{C,20}, P_{C,21} \in [0, 1] \end{aligned} \quad (11)$$

where $\pi_{D,0}$ is the marginal steady-state probability of the decompressor for state FC_0 , π_{H,s_H} is the steady-state probability of the channel in state s_H , and π_{C,s_C} is the steady-state probability of the compressor in state s_C .

D. TB-aware PDCP-level Timer-based Compressor

Practical ROHC timers are defined at the PDCP-level [3], [4], which means that a TB-level header control might require significant implementation changes. To address this difficulty, we introduce the notion of TB-aware PDCP-level Timer-based (P-timer) ROHC compressor, which has lower-layer knowledge while still operating on a packet-by-packet basis.

This header compression system consists of a deterministic timer-based compressor with timers that change packet-to-packet, a TB-level channel with CQI, and a PDCP-level decompressor with WLSB decompression at packet level. Such

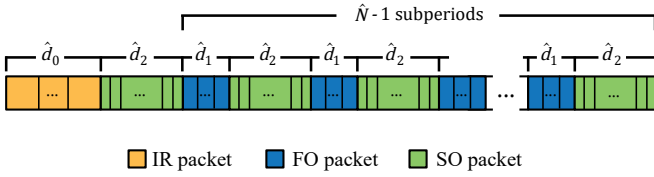


Fig. 4. PDCP-level timer-based structure. This periodic structure consists of $\hat{N}$ subperiods, the first having $\hat{d}_0$ IR packets followed by $\hat{d}_2$ SO packets, and then $\hat{N} - 1$ subperiods of $\hat{d}_1$ FO packets followed by $\hat{d}_2$ SO packets.

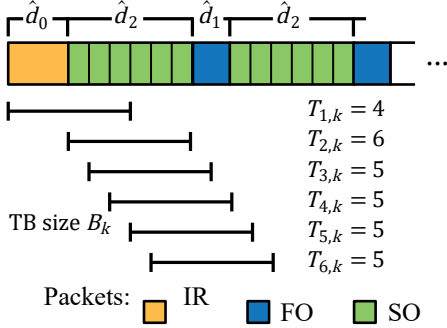


Fig. 5. Variable number of PDCP packets in a TB depending of TB size and starting position within the structure. The average of all possible arrangements can be used to compute the asymptotic efficiency of the P-timer compressor.

a system is difficult to analyze theoretically, and an exact Markov model of the system may be intractable for large timer durations. Moreover, even an approximate Markovian approach is not straightforward, as the state transitions of compressor, channel and decompressor happen in different domains and timeframes without direct models for their interaction. Therefore, we will define a slightly different framework that will allow us to optimize the P-timer compressor.

Consider the compressor with the deterministic structure of Fig. 4, where the notation has been slightly changed with respect to the TB-timer compressor in order to describe similar concepts at the packet-level, i.e., $\hat{d}_i, i \in \{0, 1, 2\}$ is the average duration of consecutive packets with IR, FO and SO headers, and $\hat{N}$ is the number of subperiods within one timer cycle. Note that for a given TB size, to fit as many whole PDCP packets as possible in the TB, there exist different packet arrangements which generate a different numbers of packets in the TB depending on the starting packet for that particular time, as shown in Fig. 5. Without the header control, we can only obtain the average number of packets per TB in state s_H , denoted as n_{s_H} , which we will use for approximation.

Let $R = \hat{d}_0 + (\hat{N} - 1)\hat{d}_1 + \hat{N}\hat{d}_2$ be the total number of packets in a timer cycle and $R_L = \hat{d}_0 L_0 + (\hat{N} - 1)\hat{d}_1 L_1 + \hat{N}\hat{d}_2 L_2$ be the total number of bytes in a timer cycle. Then there will be R different packet arrangements in the TB. For simplicity, let us initially assume that the TB size is smaller than one full timer cycle. Thus, the number of packets from a starting point t_0 and a given CQI s_H (with its corresponding

TB size B_{s_H}) can be obtained as the variable T_{t_0, s_H} in

$$\sum_{t=t_0}^{T_{t_0, s_H}} L_C[t] \leq B_{s_H} \quad \forall s_H, \forall t_0 = 1 \dots R \quad (12)$$

where $L_C : \mathbb{N} \rightarrow \{L_0, L_1, L_2\}$ is given by

$$L_C[t] = \begin{cases} L_0 & t' \leq \hat{d}_0 \\ L_1 & \text{mod}(t' - \hat{d}_0 - \hat{d}_2, \hat{d}_1 + \hat{d}_2) < \hat{d}_1 \\ & \text{and } t' > \hat{d}_0 + \hat{d}_2 \\ L_2 & \text{otherwise} \end{cases} \quad (13)$$

where $t' = \text{mod}(t - 1, R) + 1$ indicates the packet position within one timer period, with t being the packet position in the whole packet stream.

However, it is also possible that a large TB can span multiple timer cycles, namely a number of whole cycles and a partial one. Thus, we transform (12) to account for the number of whole timer cycles, since the partial cycle is the only part that can vary in number of packets aggregated in the TB:

$$\sum_{t=t_0}^{Q_{t_0, s_H}} L_C[t] \leq B_{s_H} - R_L \left\lfloor \frac{B_{s_H}}{R_L} \right\rfloor \quad \forall s_H, \forall t_0 = 1 \dots R$$

$$T_{t_0, s_H} = Q_{t_0, s_H} + R \left\lfloor \frac{B_{s_H}}{R_L} \right\rfloor \quad (14)$$

As a result, we have

$$n_{s_H} = \mathbb{E}_t\{T_{t, s_H}\} = \frac{1}{R} \sum_{t=1}^R T_{t, s_H} \quad \forall s_H \quad (15)$$

We continue to use the same TB-level FSMC model for the dynamic channel. For the decompressor, one would like to consider every possible packet arrangement and how many packets can be recovered in a successfully received TB, namely, all the packets in the TB when the decompressor is in FC_0 state, and otherwise only those after the first IR or FO packet within that particular TB. However, such an approach finds a very large state space very rapidly. Instead, we will approximate the decompressor behavior with the same TB-level structure used for the TB-timer compressor, such that the decompressor status depend on the existence of IR or FO packets within the TB instead of the leading header of the TB. This procedure will only provide an approximate but not an exact efficiency value, as the packets that might be lost before a context update in a no-context state are still accounted. Moreover, there might be more than one IR/FO packet in a TB, depending on the selection of timer parameters.

Let $\hat{s}_C = 0, 1, 2$ respectively denote the event of having at least one IR header in the TB, at least one FO header without IR, and only SO headers. The state transition probabilities of the channel-decompressor Markov system is given by

$$p(\hat{s}'|\hat{s}) = p(s'_H|s_H)p(\hat{s}'_D|\hat{s}_D, s'_H) \quad (16)$$

where $p(s'_H|s_H)$ is defined by $\mathbf{P}_H$, and

$$p(\hat{s}'_D|\hat{s}_D, s'_H) = \sum_{i=0}^1 \sum_{a=0}^2 p(\hat{s}'_D|\hat{s}_D, s_T = i, \hat{s}_C = a) \cdot p(s_T = i|s'_H)p(\hat{s}_C = a|s'_H) \quad (17)$$

where $p(s_T = 1|s'_H) = \rho_{s'_H}$, $p(\hat{s}'_D|\hat{s}_D, s_T = i, \hat{s}_C) \in \{0, 1\}$ is defined by the FSM representation of the decompressor [13, Fig. 2], and $p(\hat{s}_C = a|s'_H)$ is given by the fraction of packet header arrangements with IR, FO without IR, or only SO, for each TB size.

Therefore, and again using total probability law with steady-state probabilities, we can define the approximated asymptotic efficiency of the P-timer compressor as

$$\begin{aligned} \max \quad \hat{\eta} &= \frac{\hat{\pi}_{D,0} L_p \sum_{s_H=0}^{K-1} \pi_{H,s_H} n_{s_H}}{\sum_{s_H=0}^{K-1} \pi_{H,s_H} B_{s_H}} \\ \text{s.t.} \quad \hat{d}_0, \hat{d}_1, \hat{d}_2, \hat{N} &\in \mathbb{N}, \hat{d}_0, \hat{N} \geq 1 \end{aligned} \quad (18)$$

where $\hat{\pi}_{D,0}$ is the marginal steady-state distribution of the decompressor for state FC_0 . Note that this optimization problem is neither convex nor well structured, hence we use grid-search solutions.

IV. SIMULATION RESULTS

In this section, we present simulation results to illustrate the performance advantage of the proposed ROHC compressors compared to a POMDP compressor with link adaptation.

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 given by the number of CQI indexes in LTE and 5G NR [16, Sec. 7.1.7], [17, Sec 5.1.3]. The selected header and payload lengths are described in [11, Sec. VI-A], whereas the selected WLSB parameter W follows what was used in [7].

We implement a grid-search optimizer for both the TB-timer and P-timer ROHC compressors and SARSOP [18] for the POMDP compressor. Additionally, for every compressor design, we apply Monte Carlo simulation to test its performance. In its simulator, the P-timer compressor chooses the packet headers before the packets are aggregated into TBs of varying sizes that depend on channel state (CQI), with no knowledge of the packet header types. In the case of the TB-timer, the compressor selects the leading header of the TB as intended, before filling the rest of the TB with SO packets. The decompressor for both compressors applies WLSB to recover headers at the PDCP packet level. Therefore, in either NC or SC state, it can fully decode IR/FO-leading TBs transmitted by the TB-timer compressor. However, the decompressor can only decompress headers for packets that follow the first IR/FO header in the TB sent by the P-timer compressor, and the preceding packets are lost.

B. Performance Results

We compare the simulation performance of the optimized ROHC compressor designs (POMDP, TB-timer and P-timer) for different levels of channel SNR in Fig. 6 and Fig. 7 for two different time periods of $n = 50$ and $n = 1000$, respectively, to illustrate both the transient and the steady-state efficiency. In both cases, the TB-timer compressor clearly

TABLE I
DEFAULT SIMULATION SETTINGS FOR OUR ROHC DESIGN TESTS.

System and channel	LTE RMC R.3, no HARQ (RV=0), Single antenna, PRBs = 50, $f_D = 5$ Hz Average SNR = [0, 3, 6] dB
Header/Payload lengths	$H_0 = 59$, $H_1 = 15$, $H_2 = 1$, $L_p = 20$ bytes
WLSB	$W = 5$ ($W_{TB} = 0$)
POMDP Compressor	$P_{FA} = P_{MD} = 0.1$, $\mathbf{E}_H = \mathbf{I}_K$

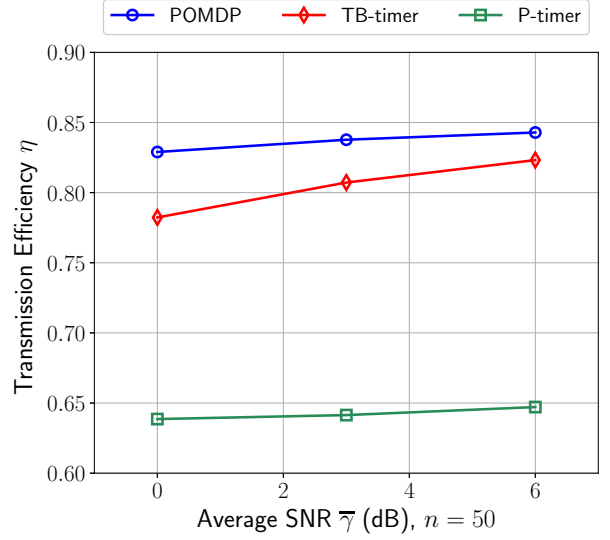


Fig. 6. The empirical efficiency η versus the average SNR of the channel, in transient state, for the POMDP, TB-timer and P-timer compressors.

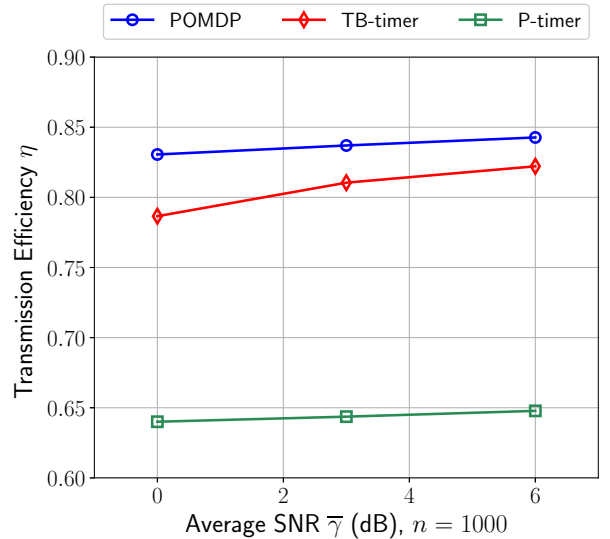


Fig. 7. The empirical efficiency η versus the average SNR of the channel, in steady-state, for the POMDP, TB-timer and P-timer compressors.

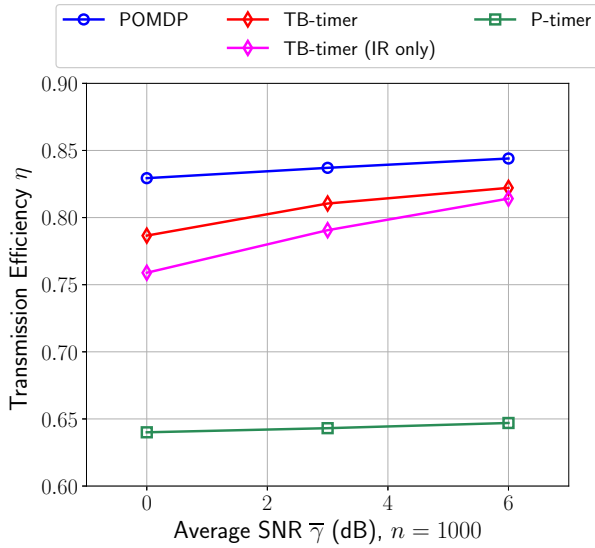


Fig. 8. The empirical efficiency η versus the average SNR of the channel, in steady-state, for the POMDP, TB-timer and P-timer compressors.

outperforms the P-timer compressor by as much as 18% under various channel quality conditions. Additionally, the TB-timer compressor efficiency is close to the optimized POMDP compressor with a maximum difference of less than 5%. This is noteworthy as the TB-level timer-based design does not depend on situational policy look-up derivations, reducing computational complexity, and thus interesting for practical implementations.

In Fig. 8 we compare the optimized TB-timer compressor with a TB-timer compressor that only selects IR-leading TBs, which would be the least-efficient TB-timer compressor design. The results show a maximum difference of 3% between the two TB-timer compressors. This comparison implies that the main contribution to efficiency gain with respect to the PDCP-level P-timer compressor arises from the TB header control strategy, again highlighting it as an attractive route for improving ROHC systems.

V. CONCLUSION

This work presents new design frameworks for U-mode ROHC compressor that takes advantage of lower-layer information in the header decision process for improving transmission efficiency. We propose a method to optimize timer-based parameters for either transport block or packet level implementations. These compressors incorporate PHY/MAC layer characteristics, such as the physical channel and link adaptation, among others. The compressors select the header compression level without direct knowledge of the ROHC decompressor status, and can easily operate with CQI report already available to the transmitting wireless unit. The new compressors can be implemented in a straightforward manner. In particular, the TB-level timer-based compressor is able to achieve near-optimum performance without incurring high complexity. Moreover, we show that the TB-based header control framework can deliver substantial efficiency gain.

REFERENCES

- [1] Federal Communications Commission (FCC), *Connecting America: 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] *5G; New Radio (NR); Physical layer procedures for control*, 3GPP Tech. Specification TS 38.213 (v15.2.0), 2018.
- [7] 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.
- [8] R. Hermenier, F. Rossetto, and M. Berlioli, “On the behavior of Robust Header Compression U-mode in channels with memory,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 3722–3732, Aug. 2013.
- [9] A. Minaburo, L. Nuaymi, K. D. Singh, and L. Toutain, “Configuration and analysis of robust header compression in UMTS,” in *Proc. IEEE 14th Personal, Indoor and Mobile Radio Commun.*, vol. 3, Sep. 2003, pp. 2402–2406.
- [10] P. Barber, “Cross Layer Design for ROHC and HARQ,” *IEEE 802.16 Broadband Wireless Access Working Group*, 2009.
- [11] 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.
- [12] C. Jiang, W. Wu, and Z. Ding, “LTE multimedia broadcast multicast service provisioning based on robust header compression,” in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, March 2017, pp. 1–6.
- [13] C. Feres, W. Wu, and Z. Ding, “A Markovian ROHC Control Mechanism Based on Transport Block Link Model in LTE Networks,” in *2018 IEEE International Conference on Communications (ICC)*, May 2018, pp. 1–7.
- [14] W. Wu and Z. Ding, “A markovian design of bi-directional robust header compression for efficient packet delivery in wireless networks,” *IEEE Transactions on Wireless Communications*, pp. 1–1, 2018.
- [15] G. Pelletier and K. Sandlund, *Robust Header Compression Version 2 (ROHCv2): Profiles for RTP, UDP, IP, ESP and UDP-Lite*, RFC 5225, Apr. 2008.
- [16] *Evolved Universal Terrestrial Radio Access (E-UTRA): Physical layer procedures*, 3GPP Tech. Specification TS 36.213 (v13.2.0), 2016.
- [17] *5G; New Radio (NR): Physical layer procedures for data*, 3GPP Tech. Specification TS 38.214 (v15.2.0), 2018.
- [18] APPL: Approximate POMDP planning toolkit. [Online]. Available: <http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/>
- [19] A. R. Cassandra. pomdp-solve v5.4. [Online]. Available: <http://www.pomdp.org/code/>
- [20] E. Patrick. POMDPy. [Online]. Available: <http://pemami4911.github.io/POMDPy/>
- [21] 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/han-nakur/dokuwiki/doku.php?id=wiki:tapir>