To Transmit or Not to Transmit: Controlling Communications in the Mobile IoT Domain

K. PANAGIDI, Department of Informatics & Telecommunications, University of Athens
C. ANAGNOSTOPOULOS, School of Computing Science, University of Glasgow
A. CHALVATZARAS, Department of Informatics & Telecommunications, University of Athens
S. HADJIEFTHYMIADES, Department of Informatics & Telecommunications, University of Athens

The Mobile IoT domain has been significantly expanded with the proliferation of drones and unmanned robotic devices. In this new landscape, the communication between the resource-constrained device and the fixed infrastructure is similarly expanded to include new messages of varying importance, control, and monitoring. To efficiently and effectively control the exchange of such messages subject to the stochastic nature of the underlying wireless network, we design a time-optimized, dynamic, and distributed decision making mechanism based on the principles of the Optimal Stopping and Change Detection theories. The findings from our experimentation platform are promising and solidly supportive to a vast spectrum of real-time and latency-sensitive applications with quality of service requirements in mobile IoT environments.

CCS Concepts: • **Computer systems organization → Embedded systems**; *Redundancy*; Robotics; • **Networks →** Network reliability.

Additional Key Words and Phrases: Real-time Decision Making, Mobile IoT, Optimal Stopping Theory, Change-point Detection, Unmanned Vehicles

ACM Reference Format:
K. Panagidi, C. Anagnostopoulos, A. Chalvatzaras, and S. Hadjiefthymiades. 2019. To Transmit or Not to Transmit: Controlling Communications in the Mobile IoT Domain. 1, 1 (October 2019), 24 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the last decade, we have been witnessing significant advancements and evolution of the Internet of Things (IoT). Going a step further to the IoT infrastructure, resource-constrained nodes are enhanced with mobility capabilities forming the Mobile IoT (MIoT) networks; noticeably, huge growth has been witnessed in the Unmanned Vehicles research area. We can consider a *drone* as a mobile computing and sensing node deployed to different locations tailored to specific tasks. The fundamental features that 'transform' Unmanned Vehicles to popular mobile IoT nodes are the ability to autonomously make decisions (i.e., without human intervention), the capability of carrying additional application-specific payloads, the endurance, capability of re-programmability,

Authors' addresses: K. Panagidi, kakiap@di.uoa.gr, Department of Informatics & Telecommunications, University of Athens, Panepistimioupolis, Ilissia, Athens, 15784; C. Anagnostopoulos, School of Computing Science, University of Glasgow, Clasgow, G12 8QQ, christos.anagnostopoulos@glasgow.ac.uk; A. Chalvatzaras, Department of Informatics & Telecommunications, University of Athens, Panepistimioupolis, Ilissia, Athens, 15784, achalv@di.uoa.gr; S. Hadjiefthymiades, Department of Informatics & Telecommunications, University of Athens, Panepistimioupolis, Ilissia, Athens, 15784, shadj@di.uoa.gr.

XXXX-XXXX/2019/10-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

and capacity to stream locally sensed/captured multimedia content. As Unmanned Vehicles become more advanced in terms of computational capabilities, they are expected to present greater value in application cases of e.g., environmental surveillance and monitoring, and supporting crisis management activities. For instance, consider the use case where drones equipped with video camera and various sensors, like air-quality, humidity and temperature, are programmed to cruise over forests and spot fires at an early stage.

The ultimate target of an Unmanned Vehicle, also coined as UxV, where 'x' can stand for either 'A' aerial, or 'G' ground, or 'S' surface vehicle, is the successful execution of a pre-programmed mission. A mission is often described as a trajectory with specific way-points that the UxV is tasked to approach and collect various measurements, e.g., from on-board sensors, or capture images or video, e.g., from on-board cameras. The way-points along with the various commands are determined from a control unit, i.e., a Ground Control Station (GCS). A GCS is a remote coordinator (master) node responsible for contextual data acquisition and real-time control and monitoring of the progress of the UxVs missions. The communication between UxV and GCS is realised in a wireless manner. The UxVs themselves can be either involved in a mission as single/individual units or as groups, i.e., swarm of UxVs. A swarm of UxVs forms a remote sensing system and can be treated as Mobile Wireless Sensor Network (MWSN) of highly dynamic topology. More importantly, the on-board computing & sensing elements of the UxVs enhance the in-network embedded intelligence of the swarm. This allows complex local computational and analytics tasks to be realized in a highly distributed fashion, thus, balancing computational load across the infrastructure and render communications much more energy efficient. In this MIoT environment of UxV-driven distributed computing, we are facing the following research and technical challenges:

Challenge 1: Real-time Monitoring. Real-time surveillance and monitoring applications, e.g., detection of forest fires, require control messages to be delivered from a *swarm* of UxVs to the GCS with the *minimal delay and high accuracy*. These missions typically involve rural areas, where the network connectivity is expected to be poor [12]. Moreover, radio paths between the UxVs and GCS are anticipated to be obstructed, overloaded or to suffer from high packet loss rate. It is challenging to predict these network variations in these environments. Hence, it is deemed crucial, during a mission, an UxV to autonomously decide *when to pause* telemetry/control measurements that are not currently prioritized as 'important' and save network resources.

Challenge 2: Secure UxV Control & Actuation. The connectivity among UxVs and GCS needs to take into consideration the *mobility* factor. This factor adds up a new degree of freedom to their operation, since the GCS sends control commands to UxVs while UxVs are moving for further local actuation. The control messages and their acknowledgements must be securely delivered in order to guarantee safe and successful missions. The usual approach to emergency cases, when a UxV loses the connection to GCS, is that the UxV returns to its initial position abandoning the mission. This means that the mission is cancelled, even if the UxV could be really close to the mission's end or objective leading to significant waste of time and resources.

In this work, we cope with the above-mentioned challenges by proposing an on-line stochastic-driven decision making scheme that leverages the transmission functionality of UxVs and GCS by being adaptive to changes in network quality. This is designed and developed by our novel suppression control of telemetry and control messages model based on the principles of the Optimal Stopping Theory (OST). Our time-optimized control mechanism achieves the *optimal delivery* of critical information from UxVs to GCS and vice-versa. Our rationale is that should the network be performing properly, then the transmission control can be 'relaxed' to exploit the available resources in the resource-constrained UxV. Our model introduces two sequential optimal stopping time decision making mechanisms based on the Change Detection theory and an application-specific discounted reward process.

We consider the case where a UxV operator desires to execute a mission and consider the setting where two main components are provided: a GCS and an Unmanned Ground Vehicle (UGV). The mission instructions could be consolidated in a domain-specific script, e.g., the mission scripts compiled through our experimentation platform for UxVs RAWFIE [17]. The RAWFIE[1] platform is briefly presented in Section 4.3. The mission script defined by the operator includes (among others) the UxV trajectory way-points in the field area to control the device in space and time and the sensing components involved (sensors) to collect in-field measurements. The main goal of the two components is the monitoring of an area to detect fire based on camera stream and on-board environmental sensors. This use case was also conducted during the RAWFIE project lifetime.

The baseline solution/establishment for the UGV's mission is as follows: The GCS sends specific commands (directives) to the UGV as indicated in an experimentation script, e.g., "Go-to-Point", "Pause" on a specific point, or "Abort" the mission and return home (RTL). The UGV sends sensor measurements streams, e.g., temperature, humidity, video and its geo-spatial position (GPS) to GCS with a predefined frequency. Both GCS and UGV have as a goal the successful completion of the monitoring of the area. Both UGV and GCS monitor the *quality of the network*. The quality of the network can be classified as proposed in [14] and is crucial for the mission because significant commands (down-link from GCS to UGV) or sensor values/measurements (up-link from UGV to GCS) can be occasionally lost due to the stochastic network behavior.

We propose a real-time control mechanism to *adapt* to changes in network quality by dynamically *pausing* control telemetry and control messages based on *optimal sequential decision making rules*. This is expected to ensure the trouble-free delivery of critical information subject to the dynamic network status that UxVs encounter while dispatching a certain mission.

Remark 1. *Overall, our scheme can be applied in all cases where connections are competing for stochastically varying network resource and optimally manage their relative priorities.*

This paper is organized as follows: In Section 2, we present the related work, while in Section 3, we present the preliminaries for our problem formulation, the proposed optimized information flow model and our two optimal stopping problem solutions. Section 4 presents our comprehensive experiments with real UxV settings, where our mechanisms performances are followed by the conclusions in Section 5.

2 RELATED WORK & CONTRIBUTION

2.1 Related Work

The challenge of optimizing contextual information flow delivery among UxVs is non trivial given the network circumstances and status. To our knowledge, there is no prior holistic work addressing the problem of time-optimized information flow. In the literature, research has been extensively focused on message-routing protocol employed on UxVs. Opportunistic networks have been proposed as long as they are capable of maintaining efficient operation in a wide range of network density and mobility conditions [26],[19]. By classifying the diversity of topological conditions in networking environments, one end of the spectrum corresponds to almost static dense topologies. In this case, conventional topology-based protocols [20] function best by using node labels / identities. As the nodal density decreases and / or the mobility increases, and up to a point where the connectivity status between pairs of nodes remains stable, position-based families of protocols [19], [11] become more suitable.

Additionally, in networks of low nodal density, intense mobility becomes a prerequisite for the creation of contact opportunities. For such topologies, protocols based on the 'carry' action [26],

[1] http://www.rawfie.eu/about

Table 1. Nomenclature

Notation	Description
DMP	Decision Making Process
TOCP	Time-Optimized Change Point DMP
DRP	Optimal Discounted Reward DMP
QNI	Quality Network Indicator
$p(x_n, f)$	Probability Density Function with parametric density f
f_i	Normal distribution $\mathcal{N}(\mu_i, \sigma_i)$
H_0	No-Change-point Hypothesis
H_1	Change-point Hypothesis
FAR	False Alarm Rate
N_d	Detection Change time
α	Detection threshold: $0 \leq \alpha \leq FAR$
$\gamma \in [0, 1]$	Discounted factor at DRP policy
$r(\gamma, N) \in \{1, \ldots, N\}$	Stopping time in DRP policy up to time N
t^*, τ^*, r^*	Optimal stopping times in generic OST, Change-point Detection, and DRP policy, respectively
T_h	Maximum horizon where a UxV can be paused
$L_x(n)$	Log-likelihood ratio at time n for random variable X

[15], i.e., the spatial transposition of the message due to the physical movement of the carrier node, perform efficiently. The aforementioned routing protocols have been designed to accommodate a restricted set of possible network conditions, corresponding to a particular sub-range and yield satisfactory performance only under these conditions.

Opportunistic Networking is also an open and an active field of research where OST can be applied at routing delivery protocols. A proposal for opportunistic networks (OppNet) [7] is studied in which the authors present Softwarecast as a general delivery scheme for group communications based on mobile code. This software code and a delivery state is the main input to persist refined delivery-decision making methods based on OST to implement complex decisions. In [10], the authors present the Relcast, a composite routing-delivery scheme that used OST-based delivery strategies to route messages to profiles which are defined by delivery functions such as best maximum and over-the-average. If we go a step further, we define a routing delivery protocol to social OppNet like influencers' networks. The [9] refer to an OST-based solution to deliver messages in highly connected networks. However, the proposed solutions are based on metrics like low latency, while the authors in [8] proposed a solution of broadcast protocols for OppNet based on efficiency, preventing unrestrained propagation of messages.

All the proposed delivery routing protocols are based on variations of the Secretary Problem [13] like the called rank-based selection and cardinal payoffs variation of the secretary problem [6]. However a unique strategy cannot be applied to sequences with abrupt changes where each state shall be treated differently. Other research efforts are focused on delay-tolerant methodologies, where mobile sinks (e.g., data aggregation nodes) 'patrol' a number of static sensor nodes and collect data [18], [29]. Nonetheless, due to their delay-tolerant principle for data delivery, they cannot be directly applied to real-time applications like disaster management.

Methods based on the principles of dynamic stochastic optimization frameworks, like Optimal Stopping Theory, have been successfully applied to information dissemination in ad-hoc networks.

The authors in [12] add mobility into wireless network infrastructure, i.e., WiFi access points (AP) on wheels, which move to optimize user performance. The Roomba devices equipped with network interfaces move independently around areas in order to maximize the wireless capacity in this area. However, the mobile devices are moving based on a grid at the floor to predefined paths. In [25], researchers apply Optimal Stopping Theory based on Change Detection only in-network statistics. This method is only applied to pause the generation of telemetry messages. Pausing period stops when a time threshold is reached and for this period framework is agnostic to network state. Researchers' method is compared with our proposed model in Section 4.3 applied on real UxVs. Contextual data delivery mechanisms have been studied in the literature though from a different 'perspective' in mobile ad-hoc networks. The contextual data delivery mechanisms in [4], [2], [24] and [3] deal with the delivery of quality information to context-aware applications in static and mobile ad-hoc networks, respectively, assuming epidemic-based information dissemination schemes. In [4], the authors propose optimal decision making approaches on the collection of contextual data from WSNs. The mechanism in [2] is based on the probabilistic extension of the well-known Secretary Problem introduced in [13] merged with an optimal on-line stochastic optimization problem. The authors in [1] tackle the task offloading decision making problem by adopting the principles of optimal stopping theory (OST) to minimize the execution delay in a sequential decision manner. Their approach significantly minimizes the execution delay for task execution and the results are closer to the optimal solution than other deterministic offloading methods. The authors in [24] study a dynamic video encoder that detects scene changes and tunes the synthesis of Groups-of-Pictures (GoP) accordingly based on an 'Black-Jack' like application of Optimal Stopping Theory. The proposed MPEG encoder tracks the error between the sequential frames in a Group-of-Pictures (GOPs) and optimally creates GOP sizes which are content-based with the minimum waste of the resources.

2.2 Contribution

Our problem deals with poor network performance during a UxV predefined mission. The online control of UxVs mission is highly connected with two types of paths: geo-spatial and network. The union of localization and network factors concludes to safe mission with accurate data reports. It is apparent that the mobility factor adds up new complexity to the aforementioned solutions in literature that handle message forwarding or routing topologies for stationary sensor networks.

Furthermore, our framework is independent of the UxVs technologies and can be applied to different kind of UxVs (aerial, sea, ground) and to their on-board software like ROS [28] or Ardupilot [5]. Mostly in literature, the UxV solutions are targeted to problems with a specific type of UxVs. However, our work in this article does not depend on the type of UxV. Our decision making process handles the control of contextual flow in a mission based on the quality network statistics with no-prior knowledge of the environment and the category of the device, i.e., aerial, ground or surface vehicles. This real-time decision making framework is based on two Optimal Stopping Time Policies that optimally schedule context delivery (control messages and values) and deliver messages with minimum loss of packets in poor or saturated networks.

Our specific technical contribution of this work is:

(1) *A stochastic optimization mechanism for on-line network quality change detection*;
(2) *A hybrid sequential decision making mechanism for optimal control commands from the GCS to UxV based on the Optimal Stopping Theory*;
(3) *Proof of optimality of the two proposed mechanisms in UxV MIoT environments*;
(4) *Comprehensive performance evaluation, sensitivity analysis of the major parameters, and comparative assessment of the proposed mechanisms in a real-testbed UxVs platform.*

Table 2. Rules of State Transitions

Component	Network State 'Good'	Network State 'Bad'
UxV - High priority sensors	ON	ON
GCS - High Priority Messages	ON	OFF
UxV - Low priority sensors	ON	OFF

3 TIME-OPTIMIZED DECISION MAKING MODEL FOR UNMANNED VEHICLES

3.1 Rationale

The main contribution of this paper is to establish an in-network/on-device lightweight sequential Decision Making Process (DMP) that leverages the on-line derived network statistics to efficiently control the progress of a UxV mission. Each UxV is equipped with a number of sensors and at least one network interface. Our DMP is capturing network related information, e.g., packet error rate, and controls the transmission of messages on both UxV to GCS and GCS to UxV, dynamically. Fundamentally, based on the real-time captured network statistics, our DMP makes transitions during the UxV mission between two states: *active* and *passive* state as shown in Figure 1. The time duration for staying in each state and the transition from one state to another are *optimally* determined by two real-time decision making mechanisms as will be discussed in the following paragraphs.

All messages exchanged between UxV and GCS are categorized in 'high' and 'low' priority. High priority message is considered (i) the minimum necessary systemic instructions to carry out a mission and (ii) sensor data defined by the UxV operator as highly important. When the UxV/GCS is in active state then the DMP sends constantly messages for telemetry and control. In the passive state, the DMP sends only high priority messages. For instance, the position reporting from the UxV is a prerequisite for the safe execution of the mission. In this case, high-priority commands are being sent constantly. Low priority messages, e.g., temperature values captured locally from the UxV sensors, can be delayed until the network exhibits *better performance*. The message priority at the GCS is the inverse, i.e., significant messages are to be delayed in order to safely reach the UxV. The described rules of state transitions based on the network state, the UxV and GCS are shown in Table 2.

The DMP runs locally on the UxV and on the GCS, enriched with a Time-Optimized Change-Point Decision Making Process (TOCP). The TOCP is triggered when a change on network performance occurs; the TOCP is discussed extensively in Section 3.3. This will enable the UxV and the GCS to transit from the *active* state to the *passive* state. When the DMP concludes on the 'passive' state, then a Discounted Reward Decision Making Process (DRP) is activated, as will be discussed in Section 3.4. The rationale is that the DRP *sequentially* ranks the network quality measurements from the relatively *worst* to the relatively *best* and, then *optimally*, it delays its pause interval for the (stochastically) globally best network observation to resume from the pausing period as dictated by the TOCP. The pausing period has a maximum deadline, hereinafter referred to as the pausing horizon Th_{max}. This indicates the maximum time interval the UxV waits *without* receiving any command and ACK messages from the GCS. To sum up, we propose a mechanism for temporal control of the transmission of the messages to and from the UxV. This mechanism is based on a network condition model that transits from good to bad and vice versa. All these transitions are monitored and validated through our system using the principles of the change detection and optimal stopping theory.

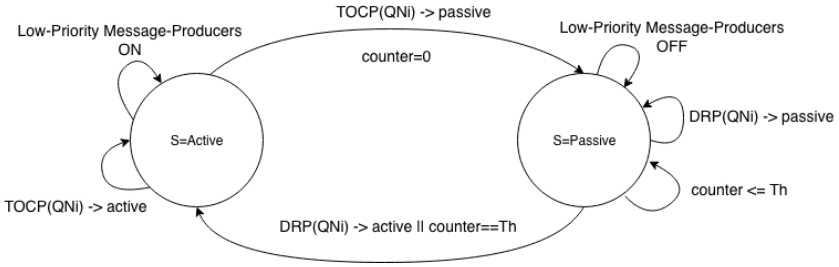


Fig. 1. UxV State Transition Model

3.2 Preliminaries in Optimal Stopping Theory & Change Detection Theory

Before elaborating on our problem formulation and the proposed time-optimized mechanisms, we provide the fundamentals and principles adopted from the OST and the change-detection theory.

3.2.1 Optimal Stopping Theory. The first studied optimal stopping problem is related with the problem of choosing a time to take a given action based on a sequentially observed random variables in order to maximize an expected payoff. In addition our stopping time problem has a finite horizon, i.e., there is an upper bound on the number of stages at which we may stop.

Let $\mathbb{F}_n$ be defined as the σ-algebra generated by the random variables $Y_1, Y_2, \cdots, Y_n$ in a probability space $(\Omega, \mathbb{F}, P)$. We envisage $\mathbb{F}_n$ as the filtration (information) observed up to (discrete) time instance n by collecting the realization values of the random variables up to n. For instance, in our context $Y_1, Y_2, \cdots, Y_n$ are considered the observed Quality of Network Indicator (QNI) values in discrete timesteps $t = 1, \ldots n$. A *stopping rule* or *stopping time* is defined as the random variable τ with realization values in a set of natural numbers such that $\{\tau = n \in \mathbb{F}_n\}$ for $n = 1, 2, \ldots$ and probability $P(\tau < \infty) = 1$. We denote with $\mathbb{M}(n, N)$ the class of all stopping rules τ in which $P(n \leq \tau \leq N) = 1$ for any $n = 1, 2, \ldots$ and $N > 0$. The real-valued pay off function is then defined as the mapping $W : \mathbb{R} \rightarrow \mathbb{R}$ being a Borel measurable function which values $W(y)$ interpret the pay off of a decision maker when it stops the Markov chain $(Y_n, \mathbb{F}_n)$ at the state $y \in \mathbb{R}$. In our case, the reward can be defined as the selection of the best network metric (QNI value) reached so far.

Assume now that for a given state y and for a given stopping rule τ, the expectation of the reward (pay-off) function is $\mathbb{E}[W(Y_\tau)|Y_1 = y]$ exists. Then, the expected pay off $\mathbb{E}[W(Y_\tau)|Y_1 = y]$ corresponding to a chosen stopping rule τ exists for all states $y \in \mathbb{R}$, which refers to the *reward value* of the stopping problem. Based on the *principles of optimality* the *reward value* $V_N(y)$ is the supremum of the expected pay off of *all* the stopping rules belonging to $\mathbb{M}(1, N)$, i.e.,

$$V_N(y) \quad = \quad \sup_{\tau \in \mathbb{M}(1,N)} \mathbb{E}[W(Y_\tau)|Y_1 = y], \tag{1}$$

where the supremum is taken for all stopping rules $\tau \in \mathbb{M}(1, N)$ for which the expectation $\mathbb{E}[W(Y_\tau)|Y_1 = y]$ exists for all $y \in \mathbb{R}$. Based on the optimal value $V_N(y)$, where the supremum in (1) is attained, the *optimal stopping rule* $t^* \in \mathbb{M}(1, N)$ should satisfy the condition:

$$V_N(y) \quad = \quad \mathbb{E}[W(Y_t^*)|Y_1 = y], \forall y \in \mathbb{R}. \tag{2}$$

It is then clear that the optimal value $V_N(y)$ is the maximum possible excepted reward to be obtained observing the random variables $Y_1, \ldots, Y_N$ up to the N-th observation.

Consider also that the expectations $\mathbb{E}[W(Y_\tau)|Y_1 = y]$ exist for all $y \in \mathbb{R}$ and, based on the principles of optimality. Let us then introduce the operator Q over the reward function $W \in \mathbb{R}$

such that:

$$QW(y) \quad = \quad \max\{W(y), \mathbb{E}[W(Y_t^*)|Y_1 = y]\}. \tag{3}$$

Then, the optimal stopping rule t^*, which attains the optimal value in (2), is estimated by the Theorem 3.1:

THEOREM 3.1. *Assume that $W \in \mathbb{R}$. Then:*

- $V_n(y) = Q^n W(y)$, $n = 1.2, \ldots$;
- $V_n(y) = \max\{W(y), \mathbb{E}[V_{n-1}(Y_1)]\}$, *where* $V_0(y) = W(y)$;
- *the optimal stopping time t_n^* is evaluated as:*

$$t_n^* \quad = \quad \min\{0 \leq k \leq n : V_{n-k}(y) = W(y)\}, \tag{4}$$

This refers to an optimal stopping rule in $\mathbb{M}(1, n)$. If $\mathbb{E}[|W(Y_k)|] < \infty$, for $k = 1, \ldots, n$, then the stopping rule t_n^* in (4) is optimal in the class $\mathbb{M}(1, n)$.

PROOF. Please refer to [13]. □

3.2.2 Change Point Detection Theory. The second category of the optimal stopping problem is the detection of a change point. Consider that we are monitoring a sequence of a sequence random variables, like values of the QNI, $\{Y_1, Y_2, \ldots Y_n\}$ with a known distribution f_0. At some point m in time, unknown to us, the distribution changes to another known distribution f_1. Our goal is to detect the change as soon as it occurs. Let $\mathbb{F}_n$, $n \geq 1$ be the σ-algebra generated by the random variables $\{Y_1, Y_2, \ldots Y_n\}$. A sequential *change point detection rule* is then derived by the stopping time τ of the observed values. The stopping time τ for the change point detection has the following characteristics:

- Average Run Length (ARL): ARL, proposed in [22], is defined as the expected number of observed values before a change decision is taken, where N_d is the detection time and f is assumed to be constant, i.e., $ARL = \mathbb{E}[N_d]$
- The Detection Delay D_n is the average detection delay corresponding to the observed $\{Y_1, Y_2, \ldots Y_n\}$ needed before a detection change occurs. Therefore, this quantity has to be as small as possible to minimize the reaction time of the algorithm.
- The False Alarm Rate *FAR* [16] is calculated as the ratio between the number of negative events wrongly categorized as changes.

In the following, we describe the two in-network/on-device optimal stopping rule mechanisms running on the UxV; the same mechanisms also run on the GCS.

3.3 Time-Optimized Change-Point Decision Making Process

3.3.1 Problem Formulation. In this section we introduce the TOCP, which reflects the behavior of the UxV being in the active state. Specifically, consider the network quality readings $x_1, x_2, \ldots, x_n$ as a discrete random signal with independent and identically distributed (i.i.d.) random variables observed sequentially in real time. Consider also that the network readings follow a probability density function $p(x_n, f_i)$. In our case, f_i expresses the normal distribution with mean value μ_i and variance σ_i. To estimate $p(x_n, f_i)$, a probability density function comparison method has been adopted to derive the *closest* distribution to our Quality Network Indicator (QNI) values.

The QNI derives from the normalization of the basic network metrics: Packet Error Rate (PER), Signal-to-Noise Ration (SNR), and the interference quality indicator (Q). The SNR is defined as the ratio of signal power to the noise power. The PER is calculated as the rate between the lost packets and the total packets sent through the network. The interference quality indicator Q is exported by an access point in the scale $[0, 100]$ and depends on the level of contention or interference,

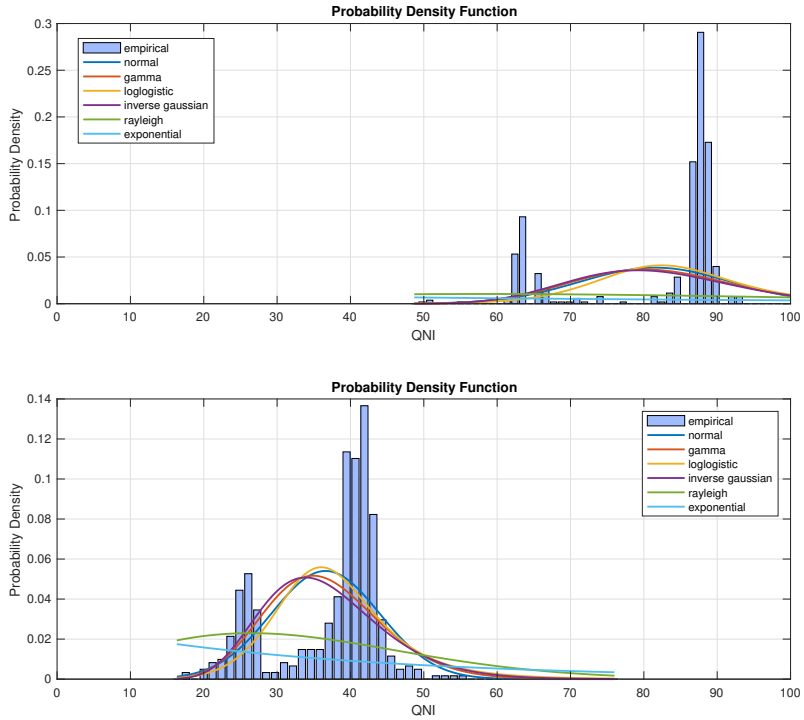


Fig. 2. (Upper) a. The probability density function f_0 and (lower) b. the f_1 model fitting for *good* and *bad* quality of QNI values, respectively.

like the bit or frame error rate, or other hardware metric. The holistic QNI at time n indicates the *quality* of the current network connectivity defined as the weighted sum of the (normalized) quality indicators:

$$QNI_n \quad = \quad a_1 P\hat{E}R_n + a_2 S\hat{N}R_n + a_3 \hat{Q}_n, \tag{5}$$

where the QNI is affine combination of PER, SNQ and Q in $[0, 100]$ such that $\sum_{i=1}^{3} a_i = 1$, $a_i \in [0, 1], \forall i$.

We consider the incoming QNI values as an adapted strong Markov process $(X_n)_{n \leftarrow 0}$ defined by the filtered probability space $p(x_n, f_0)$. The estimation of the $p(x_n, f_i)$ is based on model fitting of all the parametric probability distributions to the QNI. The output of this model fitting is shown in Figure 2a for $p(x_n, f_0)$ and Figure 2b for $p(x_n, f_1)$. The list of examined probability distributions is extensive. We are based our decisions and reasoning on the fundamental NLogL (Negative of the Log Likelihood) and the BIC (Bayesian Information Criterion) metrics. For each distribution examined, we derived the corresponding NLogL and BIC values provided in Table 3. As it is shown in Figure 2a and Figure 2b, the best distribution fitting to our experimental data is the Normal Distribution.

We further studied an abrupt change from good to bad network conditions. In this case, we performed experiments in which the network conditions changed at time m. As shown in Figure 3, before time m, the QNI follows the distribution $p(x_n, f_0)$, and after time m, the QNI follows $p(x_n, f_1)$. Under these experimental observations, the QNI distribution observed between the first sample x_0

Table 3. NLogL and BIC metrics for the probability distributions.

Examined Distribution	NLogL	BIC
Normal (N)	1876.5	3765.5
Gamma (Γ)	1904.2	3820.8
Log-logistic	1909.3	3831
Inverse Gaussian (IG)	1919.4	3851.2
Rayleigh	2366.6	4739.5
Exponential (Exp)	2700.7	54076

and the current x_k sample takes two forms, where H_0 represents No-Change-Point Hypothesis and H_1 represents the Change-Point Hypothesis:

$$p(x) = \begin{cases} \prod_{n=0}^{k} p(x_n, f_0), \text{ No-Change-Point Hypothesis } H_0; \\ \prod_{n=0}^{m-1} p(x_n, f_0) \prod_{n=m}^{k} p(x_n, f_1), \text{ Change-Point Hypothesis } H_1 \end{cases} \tag{6}$$

The challenge is to decide between the two hypotheses H_0, H_1 w.r.t. QNI, and to approximate efficiently and timely the potential change point time m. A feasible solution derived by the change-point detection theory adopts the minmax approach in [23].

Let us define the conditional expected detection delay by

$$\mathbb{E}_{H1}[(N_d - m + 1)^+ | n = 0, 1 \ldots, m - 1], \tag{7}$$

as defined in [23], where the expectation is taken under one change hypothesis H_1. The minimax performance criterion is given by its supremum taken over. Specifically, the *worst-case detection delay* is estimated as:

$$D_n(\tau) = \sup_{n \geq 1} ess\sup \mathbb{E}_k[(\tau - k + 1)^+ | \mathbb{F}_{k-1}], \tag{8}$$

with $x^+ = max\{x, 0\}$. Based on this objective, we formulate the change-point detection Problem 1:

Problem 1. *The UxV should determine an optimal change-point detection time τ that minimizes the worst-case detection delay in (8).*

3.3.2 Solution for TOCP. Let us first denote the FAR defined as [16]:

$$\text{FAR}(\tau) = \frac{1}{\mathbb{E}_\infty[\tau]}.$$

Based on our examined distribution fitting, we introduce the instantaneous log-likelihood ratio at time n by:

$$L_x(n) \quad = \quad \ln \frac{p(x(n), f_0)}{p(x(n), f_1)} = \ln \frac{\sigma_1^2}{\sigma_0^2} + \frac{(x - \mu_1)^2}{2\sigma_1^2} - \frac{(x - \mu_0)^2}{2\sigma_0^2}, \tag{9}$$

and its cumulative summation of the ratios from 0 to n:

$$S(n) = \sum_{k=0}^{n} L_x(k). \tag{10}$$

The expectation $\mathbb{E}_\infty[\tau]$ defines the expected time between false alarms. A false alarm in our case is defined when the DMP mechanism detects a change for state transition to passive, while the network quality is characterized as good. Under the Lorden criterion, our objective is to find the stopping rule that minimizes the worst-case delay subject to an upper bound on the FAR. The

decision function in our problem in a change between good and severe network conditions is shown in Figure 3*b*.

The optimal solution to (8) was determined in [21], which is provided by the Cumulative Sum (CUSUM) test [22]. A presentation of the CUSUM approach applied to our problem can be found in Appendix C and its description is shown in Algorithm 1. The optimal stopping time for detecting the change point is given by:

$$\tau^* = \min\{n \geq 1, \max_{1 \leq k \leq n} \sum_{i=k}^{n} L_x(i) \geq \alpha\} \tag{11}$$

Let the detection threshold α be chosen such that the ARL to false alarm derives $FAR \geq \alpha > 0$. Clearly, this condition is equivalent to limit the rate of false detection by a given maximum value. When $\alpha \to \infty$, the CUSUM algorithm minimizes the worst case detection delay $\mathbb{E}_{H1}[N_d]$. The value of this delay can be approximated by using Kullback-Leibler (KL) divergence. The KL captures the discrimination between the post and pre-change hypotheses and measures the *detectability* of the change, which is proved to be:

$$\mathbb{E}_{H1}[N_d] = \frac{\ln \alpha}{\log(\frac{\sigma_1}{\sigma_0}) + \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2} - \frac{1}{2}}. \tag{12}$$

See Appendix B for the derivation of the expectation $\mathbb{E}_{H1}[N_d]$.

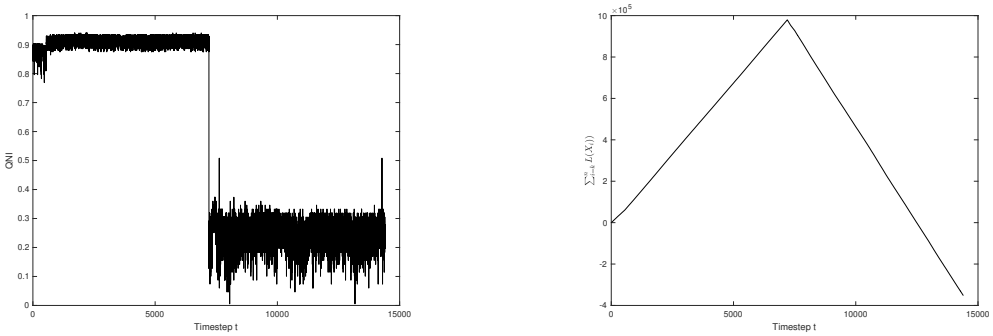


Fig. 3. The behavior of the QNI and the cumulative log-likelihood ratio corresponding to a change from a 'good' network state to a 'bad' network state.

3.4 Discounted Reward Decision Making Process

3.4.1 Problem Formulation. We propose a hybrid solution based on the change-point detection and a Discounted Reward Process (DRP) with Linear Discount Function (LDF). The reason is that the UxV cannot pause forever to send commands or to send telemetry messages. The UxV has a hard limit for sending message to GCS in order to report that it is alive and active. The same stands for GCS, i.e., the GCS cannot leave a UxV with no control messages. Therefore, the pausing period when a UxV decides whether to start again the streaming of commands can be treated as a finite horizon problem as will be described here.

It is assumed that when the pausing period starts, the UxV receives a QNI value x_k at a time instance k. The objective is to seek a stopping rule that will *maximize the probability of choosing the best (maximum) QNI value* x_k indicating the best possible network condition.

Let us define a random variable u_k, which represents the LDF reward if the kth QNI observation is chosen, that is:

$$u_k = \begin{cases} 1 - \frac{\gamma}{N}k & \text{if } x_k = \max\{x_l, l = 1, \ldots, k-1\} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

The parameter $\gamma \in [0, 1]$ denotes the *discount factor*. The discount factor γ represents the modeling abstraction where UxV focuses on selecting the best QNI value of the N received QNI values. The LDF in (13) indicates that the UxV has to report at least one QNI value observing at most N QNI values. The higher the discount factor γ is, the higher the penalty gets until a reception of a better QNI value. The UxV receives the reward u_k if the kth observation is chosen and refers to the highest QNI value among all N QNI values; otherwise u_k is zero.

Problem 2. *Given a fixed time horizon N, the UxV has to determine a optimal stopping rule $r, 1 \leq r \leq N$, which maximizes the expectation $\mathbb{E}[u_r]$.*

3.4.2 Solution for DRP. For solving Problem 2, consider first receiving the kth observation of the QNI value x_k. We can then define the random variable $z_k = j$ $(1 \leq j \leq k)$, which denotes the *relative ranking* of the QNI value x_k among the first k observations of the UxV. The assignment $z_k = 1$ means that the kth QNI value refers to the highest QNI value among the first k QNI values seen. We state then the optimal policy for a UxV w.r.t. LDF in (13) as follows in our optimal policy.

Remark 2. *Optimal Policy: There exists a time $r^*(1 \leq r^* \leq N)$ such that the UxV observes the QNI values of the first $r^* - 1$ QNI values without accepting any of them. Then for $r^* \leq k \leq N$ the UxV accepts x_k if $z_k = 1$. In case of $z_k > 1, \forall r^* \leq k < N$, or $r^* = N$, then the UxV accepts x_N, which is the last observed QNI value, with $u_N = 1 - \gamma$.*

Let $\omega_k(j)$ be the conditional expected reward of the kth observation given that $z_k = j$, that is, $\omega_k(j) = \mathbb{E}[u_k | z_k = j]$. The probability of finding the maximum QNI value x_k, i.e., $(j = 1)$, at the kth observation is:

$$P(u_k = 1 | z_k = j) = \frac{P(u_k = 1, z_k = j)}{P(z_k = j)} = \begin{cases} \frac{k}{N} & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, we have for the $\omega_k(j)$ that:

$$\omega_k(j) = \begin{cases} \frac{k}{N}(1 - \frac{\gamma}{N}k) & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

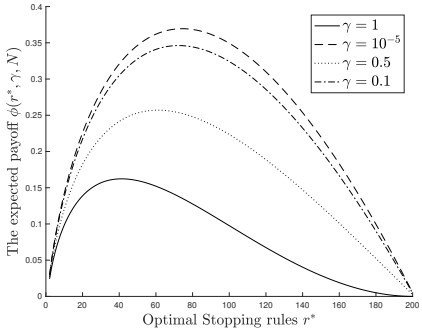
The value $\omega_k(j) = 0$ for $j \neq 1$ indicates that there is no reward if the best quality network state is not chosen.

For each $r = 1, \ldots, N$ let $\xi(r)$ denote a stopping rule, that is the first $r-1$ QNI values are observed and the next QNI value, which exceeds all of its predecessors, is accepted. If none of the first $N - 1$ QNI values is reported then the last one is reported. Then, we obtain that:

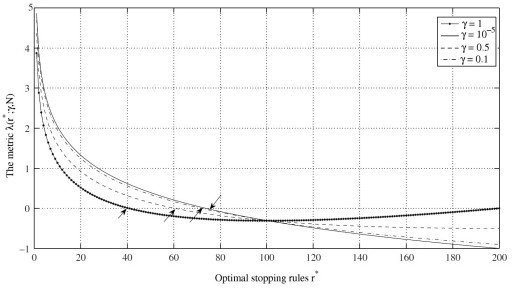
$$P(\xi(r) = k) \quad = \quad \frac{r-1}{k(k-1)}, \tag{15}$$

thus, the corresponding expected payoff $\phi(r; \gamma, N)$ w.r.t. to the reward function in (13) is

$$\phi(r; \gamma, N) = \mathbb{E}[u_{\xi(r)}] = \sum_{k=r}^{N} \omega_k(1) P(\xi(r) = k) = \frac{r-1}{N} \sum_{k=r}^{N} \left(\frac{1 - \frac{\gamma}{N}k}{k-1} \right) \tag{16}$$



(a) The expected payoff $\phi(r^*, \gamma, N)$ for different values of discount factor γ and N=200.



(b) The optimal stopping rules for different values of discount factor γ and $N = 200$. The arrows depict the earliest (optimal) stopping times r^* such that $\lambda(r^*; \gamma, N) \leq 0$.

Fig. 4. Analysis of LDF based on different values of discount factor γ.

It follows that the r^* of the proposed optimal policy that maximizes the expected payoff $\phi(r; \gamma, N)$ in (16) is the optimal stopping rule. The $\phi(r^*; \gamma, N)$ is the maximum probability of finding the best QNI value on the UxV.

Theorem 3.2. *There exists a r^* $(1 \leq r^* \leq N)$ which maximizes $\phi(r; \gamma, N)$ over $1, 2, \ldots, N$. Then, the optimal stopping rule r^* satisfies the following:*

$$r^* = r^*(\gamma, N) = \min \left\{ r \geq 1 | \lambda(r; \gamma, N) = \sum_{k=r}^{N-1} \frac{1}{k} + r \frac{2\frac{\gamma}{N}}{1 - \frac{\gamma}{N}} - \frac{1 + \gamma}{1 - \frac{\gamma}{N}} \leq 0 \right\} \tag{17}$$

Proof. See Appendix D. □

The implementation of the optimal stopping time r^* is shown in Algorithm 2. For $\gamma = 0$ and a large N, we obtain the classical optimal stopping rule $r^* = \frac{N}{e}$. Figure 4b depicts the value $\lambda(r; \gamma, N)$ and the optimal stopping rules r^* for which $\lambda(r^*; \gamma, N) \leq 0$ for different values of γ and $N = 200$. As $\gamma \to 0$ then $r^* \to \frac{N}{e}$ as illustrated in Figure 4b (for $\gamma = 10^{-5}$, $r^* = 74 \approx N/e$). The UxV reports y at observation $k \geq r^*$ for which $x_k > \max\{x_l : l = 1, \ldots, r^*\}$.

In Figure 4a we illustrate the value of the maximum probability of choosing the best QNI value $\phi(r^*; \gamma, N)$. For $\gamma = 0$ we obtain the classical secretary problem, i.e., $\phi(r^*; 0, N) \approx 1/e = 0.3678$ for large N. As the discount factor increases the maximum expected payoff decreases for large N. This indicates that we obtain a *low* likelihood (close to 0.161 for $N = 200$) in accepting the *best* QNI value once $\gamma = 1$, and this is the highest probability of achieving this. Moreover, in Figure 4b we show the optimal stopping rules for different values of discount factor γ and $N = 200$. The arrows depict the earliest (optimal) stopping times r^* such that $\lambda(r^*; \gamma, N) \leq 0$.

Remark 3. *For $0 < \gamma_1 < \gamma_2 \leq 1$ the corresponding optimal stopping rule $r_1^* > r_2^*$. This indicates that the UxV finds a QNI value earlier (stops the process earlier) when the discount factor is higher. Furthermore, as the discount factor is low then the UxV accepts a QNI value later in N; note also that the initial value of $r^* \to \frac{N}{e}$ as $\gamma \to 0$ for all N.*

Algorithm 1 TOCP-DRP Algorithm

$n \leftarrow 0$
$T_h \leftarrow$ maximum threshold interval
$r \leftarrow$ number of observations
$\alpha \leftarrow$ Change point detection threshold
$active \leftarrow TRUE$
$counter \leftarrow 0$
$(x^*, r^*) = LDS_{OST}(r, \gamma)$
while the algorithm is not stopped **do**
  if *active* **then** /* CUSUM Algorithm described in Appendix C */
    measure the current QNI x_n
    $s_n = \ln \frac{p(x(n), f_0)}{p(x(n), f_1)}$
    $S_n = \sum_{k=0}^{n} s_k$
    $G_n = S_n - \min_{1 \le k \le n}\{S_{k-1}\}$
    if $G_n > \alpha$ **then** /*A change point is detected; DRP is activated*/
      $N_d \leftarrow n$
      $\hat{n} \leftarrow \arg\min_{1 \le k \le n} S_{k-1}$
      Change occurs
      $active \leftarrow FALSE$
      Reset
    $n = n + 1$
  else
    if $n == T_h$ **then** /* maximum pausing time is reached T_h */
      $active \leftarrow TRUE$;
      break
    else$[x^*, stopped, m] = LDSF(n, r^*, x_n, x^*)$ /* invocation of DRP*/
      if $stopped == TRUE$ **then**
        $active \leftarrow TRUE$;
        break
    $n = n + 1$

Table 4. Model Parameters for the Experiments.

Parameter Names	Values
Change point detection threshold α	[0,1]
DRP discount factorγ	[1 10]
Maximum pause horizon T_h	60

4 PERFORMANCE EVALUATION

We evaluate a complete functional ground UxV that operates on two different missions, i.e., scanning search for a specific value and exhaustive scan of a certain location. We focus on the latency and the quality of the network during the mission and the impact of various parameters like mobility. We begin with a brief description of our experiment methodology.

Algorithm 2 DRP Procedures

function $LDS_{OST}(r, \gamma)$
    for $1 < n < r$ **do**
        $y(n) = LDS(n, \gamma, r)$
    return y

function $LDSF(k, r^*, x, x^*)$
    $stopped \leftarrow FALSE$
    $position \leftarrow -1$
    if $k < r^*$ **then**
        if $x > x^*$ **then**
            $x^* = x$
    else
        if $x > x^*$ **then**
            $x^* = x$
            $stopped \leftarrow TRUE$
            $position = k$
    return $x^*, stopped, position$

function LDS(x, r, γ)
    $s \leftarrow 0$
    for $x < i < \gamma$ **do**
        $s = s + \frac{1}{\gamma}$
    $y = y + r\frac{2\frac{\gamma}{N}}{1-\frac{\gamma}{N}} - \frac{1+\gamma}{1-\frac{\gamma}{N}}$
    return y

4.1 Experimental Platform & Methodology

The open-source TurtleBot device was used as ground UxV, i.e. UGV, in our experiments as shown in Figure 6. The TurtleBot uses a camera with depth sensor, i.e., XBOX Kinect for mapping purposes. ROS (Robotic Operating System) is the main operating system, which is an open-source, meta-operating system executing on a Raspberry Pi, as shown in Figure 6. UGV receives movement commands from the GCS in order to approach the given trajectory's way-points and finally reaches the objective waypoint. The UGV creates a map of the environment and, simultaneously, localizes itself in it, which is commonly known as the SLAM (Simultaneous Localization and Mapping) technology. This is also required to safely navigate within open spaces and proceed with informed decisions about the exploration targets. The Rviz [27] software was used to illustrate the mapping instance created by the UGV in Figure 8.

The communication spine between GCS and UxV is a message bus platform based on Apache Kafka, as shown in 5. The ROS publish-subscribe message pattern facilitates the interoperability with Apache Kafka, which is basically a messaging system where clients publish messages and from where consumers 'consume' them. The main advantages of the Apache Kafka are i) the high performance in delivering messages and ii) the ability to scale out by distributing the workload among different servers, therefore, supporting a cluster-based architecture. As such, it can be used for transmitting UGV measurements that will be routed from producers i.e., UxVs, to the consumers i.e., the GCS for monitoring, control, etc.

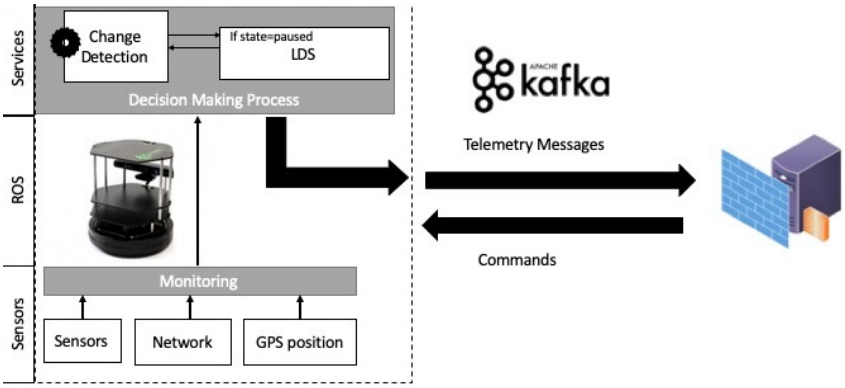


Fig. 5. The TOCP-DRP proposed architecture for the UxV Management.

(a)

(b)

Fig. 6. The Turtlebot UGV with XBOX Kinect and Raspberry Pi computing modules.

4.2 Model Parameters & Real Datasets

Prior to the real experimentation of our DMP and TOCP mechanisms, we consider a large-scale experiment generated randomly as a combination of real-life datasets. The real-life datasets were generated after multiple runs of different network conditions. We can categorize the scenarios as follows:

(1) Good dataset, experiencing no disconnections, i.e., QNI values range in $(60, 100]$;
(2) Medium dataset, indicating a saturated network where the QNI values range in $[40, 70]$;
(3) Bad dataset with several disconnections experienced, i.e., the QNI values range in $[20, 50]$.

The randomly selected blocks of all the three datasets are producing a dynamic QNI for each run of the experiment. Based on the produced dynamic QNI, we run a number of experiments in order to study the three design parameters of TOCP and DRP optimal model, i.e. α, γ, and r number of

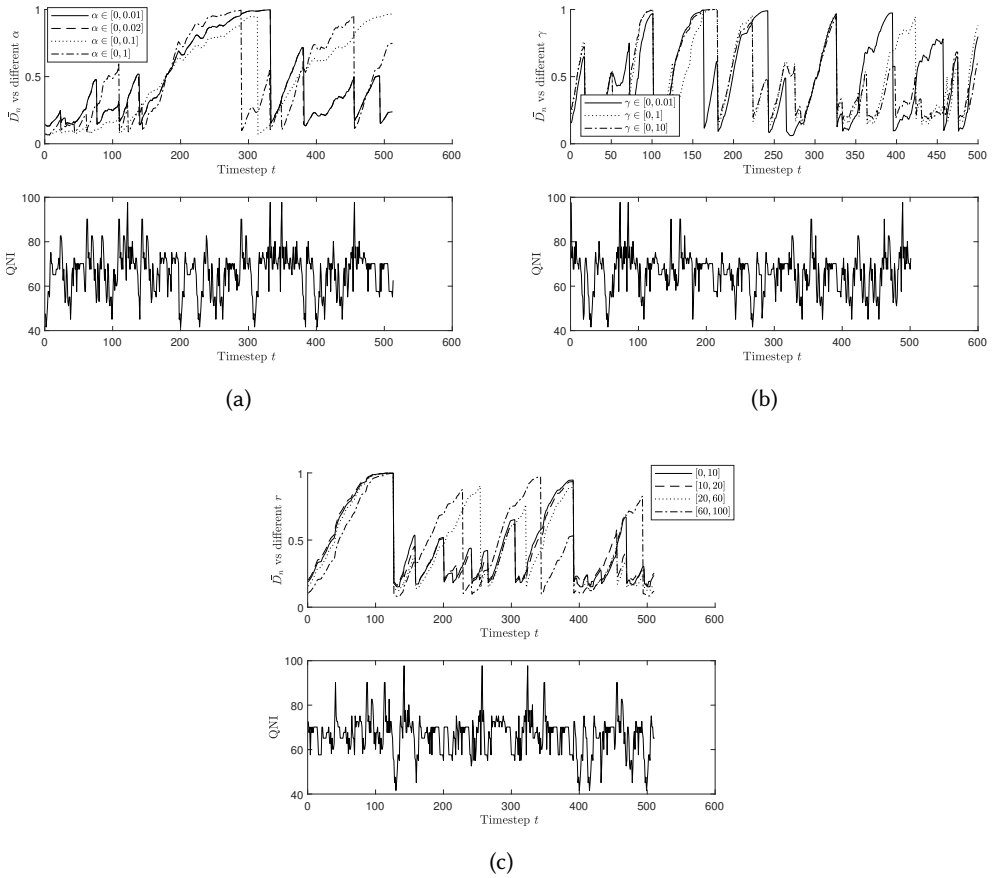


Fig. 7. The detection delay function D_n vs. (a) different α values; (b) different γ values; and (c) different observations r).

observations. We consider equal weights in equation 5 for all network parameters, i.e. $a_i = \frac{1}{3}$. We run 100 experiments with specific threshold Th_{max} and the maximum number of C_{max}. Figure 7a shows the detection delay function D_n against different α values. The D_n is more adaptive to QNI changes while α values are decreasing. The detected changes in the interval [0,0.02] are 50% more than that of $\alpha \geq 0.1$. For the DRP model, γ is a discount factor, i.e. D_n stops earlier with higher γ values as shown in Figure 7b. DRP adopts the LDF function in the 'passive' state. Therefore we can observe frequent changes from passive to the active state as expected.

We further investigate the behavior of the detection delay function D_n as r approaches infinity. As shown in Figure 7c, for small values of r, the D_n is sensitive even to small changes in network. While r is working in higher intervals, D_n is more reluctant to DRP phase. However, the probability of waiting for a large number of observations r to report a change in network quality tends to zero as shown in Figure 7c.

4.3 Experiments: Performance & Comparative Assessment

We report on the experimental evaluation of our framework and mechanisms to examine their performance. We also provide a comparative assessment with models found in the literature. The UxV and GCS are part of the Road-, Air- and Water based Future Internet Experimentation (RAWFIE) platform, which offers an experimentation framework for interconnecting numerous testbeds over which remote experimentation can be realized.

The RAWFIE platform has been developed in the context of H2020 EU-funded (FIRE+ initiative) project, which focuses on the MIoT paradigm and provides research and experimentation facilities through the ever growing domain of UxVs. The RAWFIE platform is device agnostic, promoting the experimentation under different technologies of UxVs that are equipped with different sensors, cameras and network interfaces. Any UxV is managed by a central controlling entity which is programmed per case and fully overview/drive the operation of the respective mechanisms (e.g., auto-pilots, remote controlled ground vehicles), as shown in Figure 5. The basic requirement is that each UxV shall be able to receive/send and decode/encode the incoming/outgoing messages from the testbed and deliver them to the relevant on-board component.

Our TOCP-DRP optimal mechanisms extended the functionalities of RAWFIE and can be applied to any MIoT device, i.e. UAV, UGV and USV. The used UGV in our experiments offers the convenience to make multiple repetitions of the same experiment in the campus of the University of Athens, Greece, unaffected from weather conditions and with real users.

The UGV was used in two real case applications: (1) scanning search for a specific sensor value or a detection of an event designed by a user (mission 1-M1) and (2) exhaustive scan of a room (mission 2-M2). In both missions, the user creates a path as shown in Figure 8 and the UGV should follow the way-points in order to reach the final destination. The depicted area is an amphitheater of the Department of Informatics & Telecommunications of the University of Athens and a corridor outside. During the execution of the experiments, the area is used from students and staff members that are moving around and their mobile devices are connected to the same WiFi network.

We performed 100 runs of 10 mins duration each, where each run involves sampling for more than $N = 100$ observations for every sensor integrated on UGV. The comparative assessment is based on four different policies of decision making: (i) the no-policy model, (ii) the heuristic threshold-based model, in which the transmission of messages is paused when QNI falls below a threshold, (iii) TOCP model based on [25], which applies a change detection policy triggering the 'pause' mode operation (the passive mode lasts for Th and then it is activated again) and (iv) the hybrid TOCP-DRP model applied on both UGV and GCS. The performance metrics are QNI measured, Packet Error Rate (PER), based on packets sent and packets lost, and the end-to-end message latency.

4.3.1 Expected Performance in Mission M1. Figure 9 plots the QNI performance of the four policies. We can observe that in mission M1, two areas of poor connectivity exist in time-steps [35-45] and [75-90]. The no-policy, the threshold-based policy and TOCP policy reach QNI values less than 50%, while our TOCP-DRP policy has a mean value close to 68%. In addition, for $N > 60$ the TOCP-DRP is more intolerant to network changes with mean values around [70-85].

The PER maximum values are for all the policies: $\{no-policy, threshold-based$ policy, $TOCP$, and $TOCP-DRP\}$ are $\{25, 45, 15, 10\}$, respectively, with TOCP-DRP achieving the minimum PER, i.e., we obtain up to 20% less PER compared with the other policies. The TOCP-DRP has better performance than the TOCP policy because TOCP overviews network data only in active mode and TOCP-DRP monitors QNI in both active and passive mode. The deactivation of passive mode in TOCP happens when the threshold is reached and this means that the algorithm is triggered in

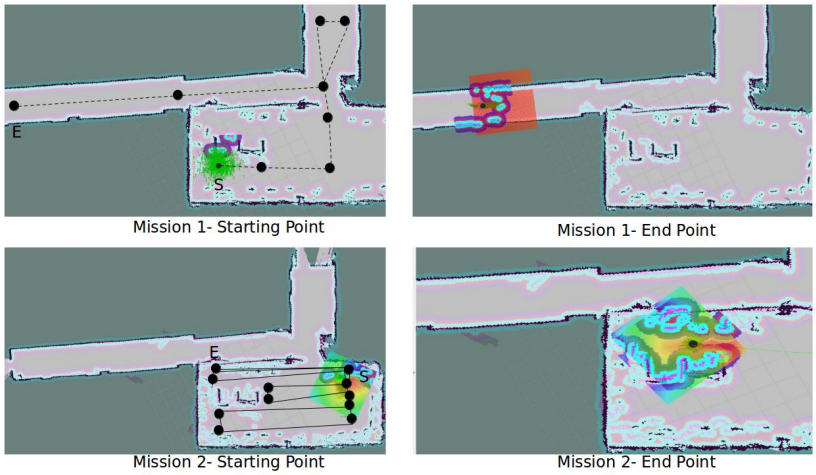


Fig. 8. Real time monitoring of the robot while executing "Mission 1-Path exploration" and "Mission 2-Scanning of an area"

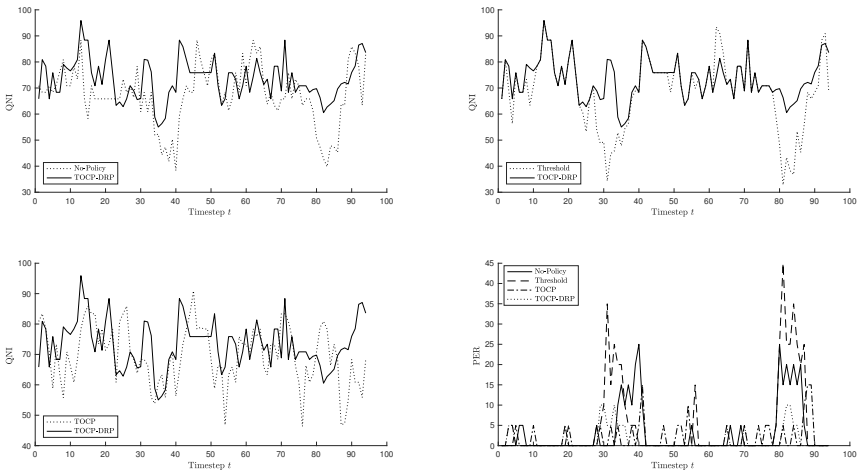


Fig. 9. The QNI for all the compared policies regarding the mission M1: Exploration of a Path.

random time-steps independently of the network status. This is the reason for observing relatively small PER values every 50 steps when the algorithm recognizes a change detection.

4.3.2 Expected Performance in Mission M2. Figure 10 shows the QNI performance of the four comparison policies for scanning missions. The M2 mission is performed indoors where areas of low connectivity and objects exist as obstacles to the UGV. The QNI has greater fluctuation in this mission relative to the M1 mission. Our TOCP-DRP mechanisms from the early beginning of mission M2, where UGV is positioned in one random corner of an amphitheater, outperforms the

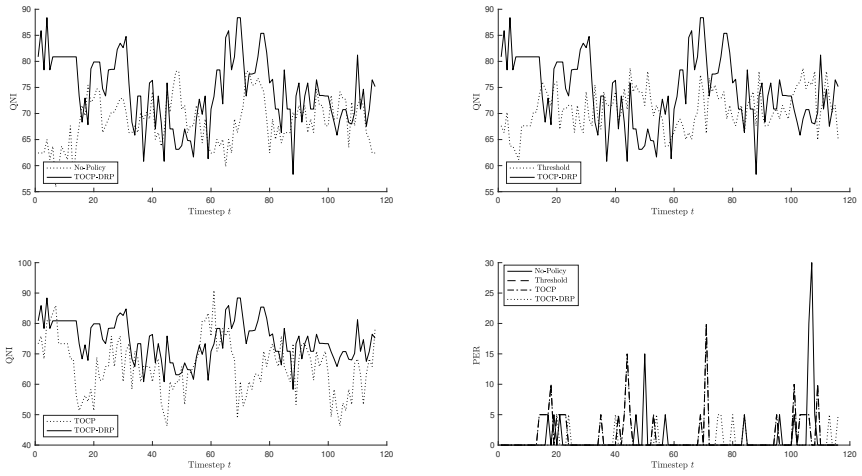


Fig. 10. The QNI for all the compared policies regarding the mission M2: Scanning of an unknown Area.

other policies. The average values of QNI for all policies: $\{no-policy, threshold-based$ policy, $TOCP$, and $TOCP-DRP\}$ are $\{68.4446, 70.8197, 65.8525, 76.3498\}$, respectively.

The performance of the PER is similar to the M1 mission. The PER is minimized in our TOCP-DRP policy, where the maximum value is 10% in observations. In the remaining policies, the PER achieve values between 20% and 30% .

4.3.3 Expected Latency in Missions M1 & M2. We plot the latency of the no-policy and our TOCP-DRP policy in Figure 11a(a) and Figure 11a(b) for the missions M1 and M2, respectively. The TOCP-DRP policy is considered more efficient than the no-policy for all the observations in both missions. In particular, in M1 we can measure 24% less end-to-end message latency compared to the original no-policy decision making of UGV. Moreover, the TOCP-DRP policy achieves systematically a message latency value which is close to 9% less of the original message latency. We can conclude that the double hybrid optimal stopping model in the two phases of the network, i.e., active and passive, based on the network assessment monitoring results to missions with low end-to-end latency and low expected PER.

5 CONCLUSIONS

We propose an in-network/on-device time-optimized decision making model of real-time control adaptive to changes of the network quality. This adaptive model dynamically pauses telemetry and control messages based on derived optimal stopping rules in order to assess in real-time the trade-off between the delivery of the messages and the network quality statistics. Our DMP policy optimally schedules critical information delivery to a back-end system. This policy uses two optimal stopping theory mechanisms based on change-detection theory and the linear discounted secretary problem. When the quality of the network significantly changes, the UxV and the GCS can decide in real-time to pause/start the transmission of telemetry in order not to overload a saturated network, or to risk to lose completely the messages. Our experimental performance evaluation and comparison assessment showed the successful delivery of messages in poor network conditions and the moderate production of messages so as not to burden an already saturated network.

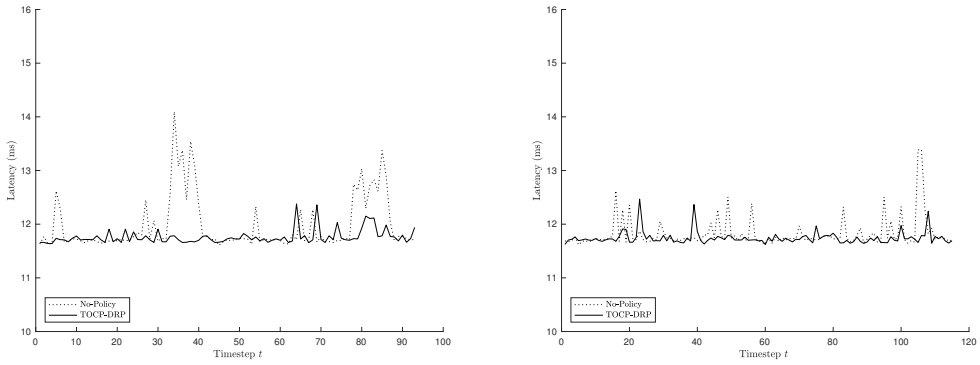


Fig. 11. The latency (ms) measured during the no-policy and the TOCP-DRP policy in mission M1 (a) and mission M2 (b).

Our future research agenda includes the adoption of our TOCP-DRP model in a swarm of UxVs in order to handle the offloading of the services / tasks, e.g., generation of telemetry, between the swarm entities. We also plan to apply our TOCP-DRP mechanism in other types of UxVs to assess in more detail the challenges in the air and/or the sea dealing with higher network quality variability.

ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Grant Agreement No 645220, project RAWFIE (Road-, Air- and Water- based Future Internet Experimentation).

REFERENCES

[1] Ibrahim Ahmed I Alghamdi, Christos Anagnostopoulos, and Dimitrios P. Pezaros. 09-13 Dec 2019. On the Optimality of Task Offloading in Mobile Edge Computing Environments. In *IEEE Global Communications Conference 2019*. Hawaii, USA.

[2] C. Anagnostopoulos and S. Hadjefthymiades. 2011. Delay-tolerant delivery quality information in ad hoc networks. *J. Parallel and Distrib. Comput.* 71(7) (2011), 974–987.

[3] C. Anagnostopoulos and S. Hadjefthymiades. Oct. 2012. Optimal Quality-aware scheduling of data consumption in mobile ad hoc networks. *J. Parallel and Distrib. Comput.* 72(10) (Oct. 2012), 1269–1279.

[4] C. Anagnostopoulos and S. Hadjiefthymiades. 2014. Advanced Principal Component-Based Compression Schemes for Wireless Sensor Networks. *ACM Trans. Sen. Netw.* 11, 1, Article 7 (July 2014), 34 pages. https://doi.org/10.1145/2629330

[5] ArduPilot Open Source Autopilot. Accessed: 2019-06-29. Open Source. http://ardupilot.org/.

[6] Neil Bearden. 2006. A new secretary problem with rank-based selection and cardinal payoffs. *Journal of Mathematical Psychology - J MATH PSYCHOL* 50 (02 2006), 58–59. https://doi.org/10.1016/j.jmp.2005.11.003

[7] Carlos Borrego, Gerard Garcia-Vandellós, and Sergi Robles. 2017. Softwarecast: A code-based delivery Manycast scheme in heterogeneous and Opportunistic Ad Hoc Networks. *Ad Hoc Networks* 55 (2017), 72–86.

[8] Carlos Borrego Iglesias, Joan Borrell, and S Robles. 2019. Efficient Broadcast in Opportunistic Networks using Optimal Stopping Theory. *Ad Hoc Networks* 88 (05 2019). https://doi.org/10.1016/j.adhoc.2019.01.001

[9] Carlos Borrego Iglesias, Joan Borrell, and S Robles. 2019. Hey, influencer! Message delivery to social central nodes in social opportunistic networks. *Computer Communications* 137 (02 2019). https://doi.org/10.1016/j.comcom.2019.02.003

[10] Carlos Borrego Iglesias, AdriÃąn SÃąnchez-Carmona, Zhiyuan Li, and S Robles. 2017. Explore and Wait: A composite routing-delivery scheme for relative Profile-casting in Opportunistic Networks. *Computer Networks* 123 (05 2017). https://doi.org/10.1016/j.comnet.2017.05.007

[11] Y. Cao and Z. Sun. 2003. Position Based Routing Algorithms for Ad Hoc Networks: A Taxonomy. *Ad Hoc Wireless Networking, Kluwer* (2003).
[12] A. Dhekne, M. Gowda, R. R. Choudhury, and S. Nelakuditi. 2018. If WiFi APs Could Move: A Measurement Study. *IEEE Transactions on Mobile Computing* 17, 10 (Oct 2018), 2293–2306. https://doi.org/10.1109/TMC.2018.2799933
[13] Thomas S. Ferguson. Accessed May, 2015. *Optimal Stopping and Applications.* Mathematics Department, UCLA.
[14] F.Fu and M. van der Schaar. 2010. Dependant optimal stopping framework for wireless multimedia transmission. In *Accoustics speech and Signal Processing (ICASSP).* IEEE, 1–6.
[15] Ai-Chun Pang Hao-Min Lin, Yu Ge and J. S. Pathmasuntharam. 2010. Performance study on delay tolerant networks in maritime communication environments. In *OCEANS 2010 IEEE-Sydney.*
[16] V. V. Veeravalli J. Unnikrishnan and S. Meyn. 2009. Least favorable distributions for robust quickest change detection. In *2009 IEEE International Symposium on Information Theory , IEEE.*
[17] Kostas Kolomvatsos, Michael Tsiroukis, and Stathes Hadjiefthymiades. 2017. An experiment description language for supporting mobile IoT applications. In *Building the Future Internet through FIRE*, Martin Serrano, Niklaos Isaris, Hans Schaffers, John Domingue, Michael Boniface, and Thanasis Korakis (Eds.). River Publishers, Gistrup, Denmark, 461–460. http://eprints.gla.ac.uk/163502/
[18] C. Konstantopoulos, G. Pantziou, D. Gavalas, A. Mpitziopoulos, and B. Mamalis. 2012. A Rendezvous-Based Approach Enabling Energy-Efficient Sensory Data Collection with Mobile Sinks. *IEEE Transactions on Parallel and Distributed Systems* 23, 5 (May 2012), 809–817.
[19] S. Mao-Y. Xiao I. Chlamtac M. Chen, V. Leung. 2009. Hybrid Geographical Routing for Flexible Energy-Delay Trade-Offs. *IEEE Trans. Veh. Technol* 58, 9 (2009), 4976–4988.
[20] A. B. McDonald. April 1997. Survey of adaptive shortest-path routing in dynamic packet-switched networks. *Technical report at the Dept of Information Science and Telecommunications* (April 1997).
[21] G. V. Moustakides. 1986. Optimal Stopping Time for Detecting Changes in Distributions. *Ann. Statist.* 14, 4 (1986), 1379–1387.
[22] E. S. Page. 1954. Continuous inspection schemes. *Biometrica* 41 (1954), 100–115.
[23] E. S. Page. 1971. Procedures for reacting to a change in distribution. *Ann. Math. Statist.* 42, 6 (1971), 1897–1908.
[24] K. Panagidi, C. Anagnostopoulos, and S. Hadjiefthymiades. 2017. Optimal Grouping-of-Pictures in IoT video streams. *Computer Communications - In press* (2017).
[25] K. Panagidi, I. Galanis, C. Anagnostopoulos, and S. Hadjiefthymiades. 2018. Time-Optimized Contextual Information Flow on Unmanned Vehicles. In *2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob).* 185–191.
[26] L. Blazevic S. Giordano, I. Stojmenovic. 2013. Routing in Delay/Disruption Tolerant Networks: A Taxonomy, Survey and Challenges. *IEEE Communications Surveys and Tutorials* 15, 2 (2013), 654–677.
[27] Rviz ROS Software. Accessed: 2019-06-29. Open Source. http://wiki.ros.org/rviz.
[28] Robot Operating System. Accessed: 2019-06-29. Open Source. https://www.ros.org/.
[29] Y. Wang, W. Peng, and Y. Tseng. 2010. Energy-Balanced Dispatch of Mobile Sensors in a Hybrid Wireless Sensor Network. *IEEE Transactions on Parallel and Distributed Systems* 21, 12 (Dec 2010), 1836–1850. https://doi.org/10.1109/TPDS.2010.56

A APPENDIX

Proof. The function $L^*(\cdot)$ of the log-likelihood ratio between $\bar{f}_0$ and $\bar{f}_1$ is continuous over the support of $\bar{f}_1$ and has an extremum. The proof is based on the first derivative test:

$$\frac{dL}{dx} = \frac{2(x-\mu_1)}{2\sigma_1^2} - \frac{2(x-\mu_0)}{2\sigma_0^2} = x\left(\frac{\sigma_0^2-\sigma_1^2}{\sigma_0^2\sigma_1^2}\right) + \frac{\mu_0\sigma_1^2-\mu_1\sigma_0^2}{\sigma_0^2\sigma_1^2}. \quad (18)$$

For $\mu_0 > \mu_1$ and $\sigma_1 > \sigma_0$, we obtain that $x^* = \frac{\mu_1\sigma_0^2-\mu_0\sigma_1^2}{\sigma_0^2-\sigma_1^2}$.

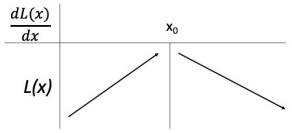


Fig. 12. Monotony analysis of L

□

B APPENDIX

Proof. The KL divergence captures the discrimination between the post and pre-change hypotheses and is a measure of the tractability of the change:

$$\mathbb{E}_{H1}[N_d] = \frac{ln\alpha}{I(p_{f0}, p_{f1})} = \frac{ln\alpha}{E_{f0}[ln(\frac{p(x_n,f0)}{p(x_n,f1)})]}, \tag{19}$$

where

$$\begin{aligned}
I(p_{f0}, p_{f1}) &= E_{f0}(ln(\frac{p(x_n,f0)}{p(x_n,f1)})) = \int [(log(f0) - log(f1))]f0dx \\
&= \int [\frac{-1}{2}ln(2\pi) - ln(\sigma_0) - \frac{-(x-\mu_0)^2}{2\sigma_0^2} + \frac{1}{2}ln(2\pi) + ln(\sigma_1) + \frac{(x-\mu_1)^2}{2\sigma_1^2}] \\
&\quad \times \frac{1}{\sqrt{2\pi\sigma_0^2}} exp[-\frac{(x-\mu_0)^2}{2\sigma_0^2}]dx \\
&= \int \{log(\frac{\sigma_1}{\sigma_0}) + \frac{1}{2}[\frac{(x-\mu_1)^2}{\sigma_1^2}] - \frac{(x-\mu_0)^2}{\sigma_0^2}\} \times \frac{1}{\sqrt{2\pi\sigma_0^2}} exp[-\frac{(x-\mu_0)^2}{2\sigma_0^2}]dx \\
&= E_0\{log(\frac{\sigma_1}{\sigma_0}) + \frac{1}{2}[\frac{(x-\mu_1)^2}{\sigma_1^2}] - \frac{(x-\mu_0)^2}{\sigma_0^2}\} \\
&= log(\frac{\sigma_1}{\sigma_0}) + \frac{1}{2\sigma_1^2}E_0\{(X-\mu_1)^2\} - \frac{1}{2\sigma_0^2}E_0\{(X-\mu_0)^2\} \\
&= log(\frac{\sigma_1}{\sigma_0}) + \frac{1}{2\sigma_1^2}[E_0\{(X-\mu_1)^2\} + 2(\mu_0-\mu_1)E_0(X-\mu_0) + (\mu_0^2-\mu_1^2)] - \frac{1}{2} \\
&= log(\frac{\sigma_1}{\sigma_0}) + \frac{\sigma_0^2 + (\mu_0-\mu_1)^2}{2\sigma_1^2} - \frac{1}{2}
\end{aligned}$$

□

C APPENDIX

In the CUSUM algorithm, we further define the generalized log-likelihood ratio G_x

$$\begin{aligned}
G_x[k] &= max_{1\le m\le k} L_x[k,m] = max_{1\le m\le k} \sum_{n=m}^{k} ln\frac{p(,x(n)f_0)}{p(,x(n)f_1)}, \\
&= S[k] - min_{1\le m\le k} S[m-1],
\end{aligned}$$

where $\hat{m}$ is defined as

$$\hat{m} = \arg\min_{1\le m\le k} S[m-1] \tag{20}$$

Equation 20 shows that the decision function $G[k]$ is the current value of the cumulative sum $S[k]$ minus its current minimum value. Equation 20 shows that the change time estimate is the time following the current minimum of the cumulative sum. Therefore, each step composing the whole algorithm relies on the same quantity: the cumulative sum $S[k]$. This explains the name of cumulative sum or CUSUM algorithm.

D APPENDIX

Proof. The expected payoff of the stopping rule r is $\phi(;\gamma,N)$. Hence, we find the first optimal stopping rule r^* for which it holds true that $\phi(r;\gamma,N)-\phi(r+1;\gamma,N)\geq 0$, to stop at r given the conditional expectation of the reward at $r+1$ after observing the relative rankings up to r. Specifically, since the conditional expectation at $r+1$ is

$$\phi(r+1;\gamma,N)=\frac{r}{N}\sum_{k=r+1}^{N}\frac{1-\frac{\gamma}{N}k}{k-1} \tag{21}$$

we can derive that:

$$\phi(r;\gamma,N) \quad = \quad \frac{r-1}{N}\sum_{k=r}^{N}\frac{1-\frac{\gamma}{N}k}{k-1}=\frac{1-\frac{\gamma}{N}r}{N}+\phi(r+1;\gamma,N)-\frac{1}{N}\sum_{k=r+1}^{N}\frac{1-\frac{\gamma}{N}k}{k-1}.$$

Hence, in order to stop at the first r, which satisfies that:

$$\phi(r;\gamma,N)-\phi(r+1;\gamma,N)\geq 0, \tag{22}$$

we obtain that:

$$(1-\frac{\gamma}{N}r)+\frac{\gamma}{N}(N-r)-(1-\frac{\gamma}{N})\sum_{k=r}^{N-1}\frac{1}{k}\geq 0 \tag{23}$$

which concludes

$$\sum_{k=r}^{N-1}\frac{1}{k}+r\frac{2\frac{\gamma}{N}}{1-\frac{\gamma}{N}}-\frac{1+\gamma}{1-\frac{\gamma}{N}}\leq 0. \tag{24}$$

Hence, the optimal stopping time r^* is obtained at the first $r\geq 1$, where the above equation turns non-positive. □