SELF-ADAPTIVE MACHINE LEARNING OPERATING SYSTEMS FOR SECURITY APPLICATIONS

Kyeong Seon Kim, Dohyun Kwon, Yejin Kim, Junhui Kim, Joongheon Kim

School of Computer Science and Engineering, Chung-Ang University, Seoul, Republic of Korea

ABSTRACT

This paper proposes a reliable and self-adaptive operating system management policy for CCTV-based security applications which controls arrival image compression rates. After receiving image sequences via CCTV cameras, the system enqueues the sequences of images and processes them for face recognition. High compression rates in CCTV-recorded images provide low recognition performance due to quantization while it is beneficial in terms of queue stability. On the other hand, low compression rates in the images provide high recognition performance while it may introduce queue overflows. Therefore, this paper designs a queue-aware self-adaptive reliable operating system management scheme which aims at face identification performance maximization while avoiding queue overflow by controlling CCTV-recorded image compression rates based on the theory of Lyapunov optimization.

Index Terms— Surveillance Monitoring, Camera Networks, Lyapunov Optimization, Stochastic Optimization

1. INTRODUCTION

Every single day, brand-new machine learning and deep learning algorithms have been investigated in many computer science research areas such as computer vision, wireless networks, embedded and operating systems, and big-data information processing [1, 2]. The proposed learning algorithms obviously present excellent performances in terms of their own objectives, i.e., maximization of recognition accuracy. However, they can be quite computationally expensive to be operated in performance-limited computer systems such as mobile smartphone devices. Therefore, in the viewpoints of systems engineers, it is crucial to think about system-level supports for computation-rich learning algorithms.

This paper designs CCTV-equipped computer vision embedded system platforms which *reliably* conducts face identification where the *reliability* is defined as the avoidance of system queue overflow due to large computation delays in data-intensive applications [3, 4, 5, 6, 7, 8].

Our considering platform is equipped with multiple CCTV cameras which individually compress recorded digital images in real-time and conducts face identification. If the image compression rate is low (which is good for identifying fast moving human faces), it introduces delays due to large number of frames. This occurs queue-overflow that is obviously system-unstable (i.e., tradeoff between stability and identification performance). Thus, a new system-support algorithm which aims at time-average face identification performance maximization subject to system reliability under the concept of Lyapunov optimization framework.

This paper is organized as follows: Section 2 and 3 present the preliminary knowledge and the proposed time-average optimization for self-adaptive CCTV-based security applications. The performance evaluation results and conclusions are presented in Section 4 and 5, respectively.

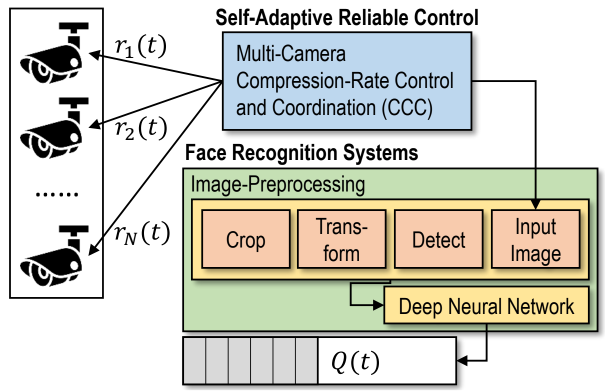

Fig. 1: Reference System Model

2. PRELIMINARIES

2.1. Reference Network Model

Our considering system model for self-adaptive face recognition operating systems is shown in Fig. 1. As illustrated in Fig. 1, multiple CCTV cameras exist for surveillance monitoring. In each camera, the recorded analog signals will be converted to digital, and then, certain amounts of compression will conduct for high-speed real-time data communications; and the individual compression ratios will be determined by *self-adaptive control for reliable systems* component.

As presented in Fig. 1, multiple CCTV cameras are connected to *self-adaptive control for reliable systems* component. If the queue-backlog is large in system queue/buffer ($Q[t]$ in Fig. 1), having less compression in incoming images via multiple cameras is harmful because it can introduce queue-overflow. Therefore, high image compression rates are desired in this case. On the other hand, if the queue is idle, having less compression can be utilized for enhancing face recognition accuracy. Notice that high compression in images definitely introduces less recognition accuracy. Finally, the *self-adaptive control for reliable systems* component determines the amounts of bitrates from all connected multiple CCTV cameras.

The actual face recognition conducted in *face recognition systems*. It consists of two stages, i.e., image-preprocessing and deep neural network. The image-preprocessing is for manipulating input image streams for learning; and it consists of four steps, i.e., input image loading, detect face region, transformation, and crop (i.e., removing backgrounds and extracting only face regions). After that, deep neural network is used for identifying the people on the images.

2.2. Time-Average Optimization with Lyapunov Drifts

The theory of stochastic network optimization aims at time-average utility optimization while achieving queue/system stability when the tradeoff relationship exists between objective function and queue stability. In the time-average stochastic optimization formulation, the concept of Lyapunov control theory is utilized for the modeling of queue stability [9]. According to the theory, it takes the minimum of the Lyapunov drift leads to the queue stability while pursuing the minimization of the time-average objective function. By taking a control action to minimize the both of time-average objective and Lyapunov drift in each unit time, time-average optimization can be obtained with the gap of $O(1/V)$ from optimality while satisfying queue stability and a time-average queue backlog bound of $O(V)$ where V is defined as a tradeoff factor between utility and stability.

More details about the theory of stochastic optimization with Lyapunov control, named to *drift-plus-penalty (DPP) algorithm*, are presented in [9]. In addition, the various applications of the theory are summarized in Section 2.3.

2.3. Related Work

As well-discussed in Section 2.2, the Lyapunov optimization based DPP algorithm which is based on stochastic network optimization (time-average utility function optimization while achieving queue stability) is scalable, and thus there exit a lot of applications for the DPP algorithm as follows:

- *Video Applications:* J. Kim, *et. al.* [10] proposed an algorithm for time-average video streaming quality maximization subject to transmission queue stability in device-to-device (D2D) video delivery. The corresponding Android software implementation is also demonstrated in [11]. J. Koo, *et. al.* [12] proposed a dynamic adaptive streaming over HTTP (DASH) algorithm for time-average video streaming quality maximization under the consideration of energy status, LTE data quota, and transmission queue stability in integrated LTE/WiFi networks.
- *Communications Applications:* M. J. Neely, *et. al.* [13] designed an energy-efficient multi-hop routing which is for time-average energy consumption minimization subject to node queue stability. Its corresponding practical implementation was presented and discussed in [14].
- *Surveillance Applications:* Y. J. Mo, *et. al.* [15] illustrated parallel machine learning systems for CCTV-base security applications with single CCTV camera. In the system, multiple artificial neural network (ANN) frameworks exist; and each ANN is with its own configurations; and it has tradeoff between complexity and performance. Therefore, the proposed algorithm adaptively selects an ANN depending on queue-backlog for time-average recognition accuracy maximization subject to CCTV queue stability. D. Kim, *et. al.* [16] illustrated face identification machine learning architectures for CCTV-base security applications with single CCTV camera. In stead of having multiple ANN frameworks, this architecture has one learning software system (based on OpenFace software library) and controls the sampling rates of CCTV camera. Therefore, the proposed algorithm dynamically controls sampling rates for time-average recognition accuracy maximization subject to CCTV queue stability.

In this paper, our proposed algorithm controls the compression rates in multiple CCTV cameras while maximizing time-average recognition accuracy maximization subject to stability for surveillance monitoring applications. Comparing to the previous single CCTV dynamic algorithms [15, 16], the proposed algorithm in this paper additionally controls the multi-camera scheduling and coordination.

3. STOCHASTIC ADAPTATION FOR SELF-ADAPTIVE CCTV SYSTEMS

3.1. Algorithm Overview

The proposed algorithm in this paper consists of two stages:

- *Stochastic Compression Rate Adaptation (Stage 1):* This stage controls bitrates from all CCTV cameras for time-average face-identification accuracy maximization subject to queue stability. More details are in Section 3.2.1.
- *Multi-Camera Scheduling and Coordination (Stage 2):* After computing the amount of time-average optimal bitrate, the compression rates of individual CCTV cameras and their corresponding coordination will be determined in this stage. More details are in Section 3.2.2.

3.2. Proposed Two-Stage Algorithm

3.2.1. Stage 1: Stochastic Compression Rate Adaptation

This section describes how we model the amount of time-average optimal bitrate using the Lyapunov optimization [9]. Based on the obtained bitrate value in this stage, the compression ratio values in individual multi-CCTV cameras are determined in the next stage (refer to Section 3.2.2).

We first model the queue dynamics as follows [9]:

$$Q(t+1) = \max\{Q(t) - \mu(t), 0\} + \lambda(t), \qquad (1)$$

where $Q(t)$, $\mu(t)$, and $\lambda(t)$ respectively denote the queue-backlog size (and $Q(0) = 0$), the number of images departing from the queue, and the number of images arriving in the queue (departure images from *face recognition systems* in Fig. 1). As shown in Fig. 1, the arrived images into *face recognition systems* is equivalent to the images which are leaving from the system because we identify the faces from the images and we are not manipulating the images at all. Therefore, the $\lambda(t)$ in (1) is equivalent to the bitrate from all connected cameras. It represents the fact that we have to compute the $\lambda(t)$ in (1) which can maximize time-average identification accuracy performance subject to stability.

We then formulate the mathematical program for maximizing the time-average face identification accuracy performance $P(\lambda(t))$ where given bitrate is $\lambda(t)$ can be presented:

$$\max : \lim_{t\to\infty} \sum\nolimits_{\tau=0}^{t-1} P(\lambda(\tau)) \qquad (2)$$

and queue stability constraint: $\lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} Q(t) < \infty$.

According to the Lyapunov optimization theory based DPP algorithm [9], this program can be re-formulated as following where $\lambda^*(t)$ is time-average optimal bitrate:

$$\arg\max\nolimits_{\lambda(t)\in\Lambda} \{V \cdot P(\lambda(t)) - Q(t)\cdot\lambda(t)\} \qquad (3)$$

where $\lambda^*(t)$ is optimal bitrate for time-average identification accuracy maximization subject to stability, Λ is the set of all possible bitrates, and V is the tradeoff coefficient between performance and stability, respectively.

Semantically, this (3) can be evaluated as follows:

- *Idle Queue Case:* Suppose that $Q(t) = 0$. Based on (3),

$$\begin{aligned}\lambda^*(t) &\leftarrow \arg\max_{\lambda(t)\in\Lambda} \{V\cdot P(\lambda(t)) - 0\cdot\lambda(t)\} && (4)\\ &= \arg\max_{\lambda(t)\in\Lambda} V\cdot P(\lambda(t)) && (5)\end{aligned}$$

  thus we have to choose $\lambda(t)$ which maximizes $P(\lambda(t))$. It is obvious that larger bitrate guarantees higher identification accuracy. Therefore, $P(\lambda_1(t)) \geq P(\lambda_2(t))$ when $\lambda_1 \geq \lambda_2$. Thus, it is finally true that $\lambda^*(t)$ will be the maximum value among the elements in the set of Λ. This is semantically true because it is beneficial to select the maximum bitrate to maximize accuracy if the queue is idle.

- *Busy Queue Case:* Suppose that $Q(t) \approx \infty$. Based on (3),

$$\begin{aligned}\lambda^*(t) &\leftarrow \arg\max_{\lambda(t)\in\Lambda} \{V\cdot P(\lambda(t)) - \infty\cdot\lambda(t)\} && (6)\\ &\approx \arg\min_{\lambda(t)\in\Lambda} \lambda(t) && (7)\end{aligned}$$

  thus we have to choose minimum $\lambda(t)$ among the elements of Λ. This is semantically true because it is beneficial to select the minimum bitrate to avoid queue overflow if the queue is almost near overflow.

Therefore, our proposed closed-form equation (3) should be computed in each unit time after observing $Q(t)$, and then it can guarantee time-average identification accuracy maximization subject to stability. Based on this nature, our proposed algorithm is self-adaptive because it can control its own bitrates automatically. In addition, this algorithm is reliable due to the fact that the self-adaptation is for maximizing its utility while achieving *stability*.

3.2.2. Stage 2: Multi-Camera Scheduling and Coordination

After computing optimal bitrate at time t, i.e., $\lambda^*(t)$, each CCTV camera computes its own compression ratio. Suppose that each CCTV camera and its corresponding compression ratio at time t are denoted by $\mathcal{C} \triangleq \{c_1, \cdots, c_N\}$ and $\mathcal{R} \triangleq \{r_1(t), \cdots, r_N(t)\}$, respectively.

At time t, the entire bitrates (without compression) from all CCTV cameras can be formulated as follows:

$$I^*(t) = \sum\nolimits_{i=1}^{N} I_i(t) \qquad (8)$$

when $I_i(t)$ is the recorded bitrate at CCTV camera $c_i \in \mathcal{C}$.

Our policy works as follows:

- If $I^*(t) \leq \lambda^*(t)$, compression is not required. Then the optimal compression ratio in each CCTV camera $c_i \in \mathcal{C}$ can be denoted as follows:

$$r_i^*(t) \leftarrow r_i(t), \forall i \in \{1, \cdots, N\} \qquad (9)$$

- If $I^*(t) > \lambda^*(t)$, proportional compression is utilized in this paper which control compression ratios depending on the image sizes in each CCTV camera, i.e.,

$$r_i^*(t) \leftarrow \frac{\lambda^*(t)}{I^*(t)} r_i(t), \forall i \in \{1, \cdots, N\} \qquad (10)$$

3.3. Pseudo-Code and Computational Complexity

The proposed algorithm in Section 3.2 (Stochastic Compression Rate Adaptation) and Section 3.3 (Multi-Camera Scheduling and Coordination) is presented in the form of pseudo-code (Algorithm 1). As shown in Algorithm 1, the Stage 1 and Stage 2 procedures are closed-form equation solving. Thus, the proposed algorithm is polynomial-time.

Algorithm 1 Proposed Two-Stage Algorithm

Initialize:
$t \leftarrow 0$
$Q(t) \leftarrow 0$
Camera Set: $\mathcal{C} = \{c_1, \cdots, c_N\}$
Camera Ratio Set: $\mathcal{R} = \{r_1(t), \cdots, r_N(t)\}$
Stochastic Compression Rate Adaptation (Stage 1):
while $t \leq T$ **do** // T: operation time
    Observe $Q(t)$
    $\mathcal{T}^* \leftarrow -\infty$
    for $\lambda(t) \in \Lambda$ **do**
        $\mathcal{T} \leftarrow V \cdot P(\lambda(t)) - Q(t) \cdot \lambda(t)$
        if $\mathcal{T} \geq \mathcal{T}^*$ **then**
            $\mathcal{T}^* \leftarrow \mathcal{T}$
            $\lambda^*(t) \leftarrow \lambda(t)$
Multi-Camera Scheduling and Coordination (Stage 2):
    $I^*(t) = \sum_{i=1}^{N} I_i(t)$
    if $I^*(t) \leq \lambda^*(t)$ **then**
        $r_i^*(t) \leftarrow r_i(t), \forall i \in \{1, \cdots, N\}$
    else
        $r_i^*(t) \leftarrow \frac{\lambda^*(t)}{I^*(t)} r_i(t), \forall i \in \{1, \cdots, N\}$

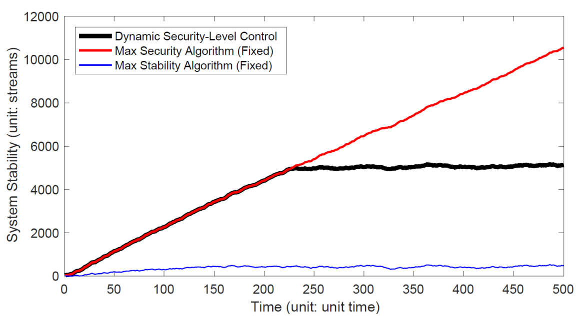

Fig. 2: Evaluation results (system stability).

4. PERFORMANCE EVALUATION

For evaluating the performance, we implemented the simulator for self-adaptive face identification frameworks with TensorFlow software library. Then, we numerically measured the tradeoff between recognition accuracy and computation time. Based on the accuracy and computation time traces via TensorFlow scripts, we simulated our platform with random event arrivals which is equivalent to the CCTV stream arrivals.

Based on the software implementation, the queue dynamics can be simulated as shown in Fig. 2. The simulation results of the proposed stochastic algorithm (denoted by *Dynamic Security-Level Control (DSLC)* were compared with the queue-backlog information of *Max Security (MaxSec)* (which is for the minimization of multi-CCTV compression ratios for maximum identification accuracy) and *Max Stability (MaxStab)* (which is fhr the maximization of multi-CCTV compression ratios for queue stability). In this simulation-based study, the proposed DSLC adaptively controls the

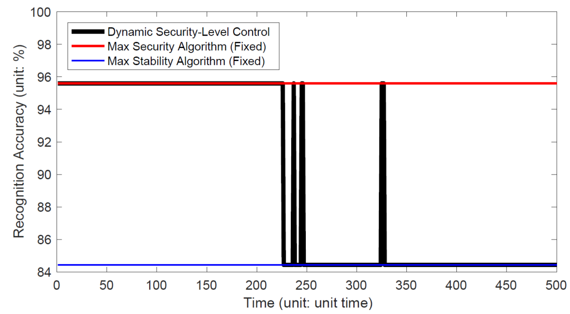

Fig. 3: Evaluation results (recognition accuracy).

compression ratios in each CCTV camera in each unit time. Fig. 2 shows that the MaxStab is extremely queue-stable because of the maximum compression ratios in all CCTV cameras for avoiding queue overflows. On the other hand, the MaxSec is not system-stable at all times because of minimum compression ratios in all CCTV cameras for achieving the highest identification accuracy. The proposed DSLC shows similar performance with MaxSec in initial times because the queue-backlog is not too much to start the adaptive control; as well as our main objective is the maximization of identification accuracy. When the unit time meets 280, the proposed DSLC starts to process bitrates from the queue for queue stability by selecting optimal bitrates (in turn, compression ratios in each CCTV cameras with proportional scheduling and coordination) based on the queue-backlog information.

In Fig. 3, the recognition accuracy adaptation can be observed. In this simulation, three frameworks are assumed thus three different accuracy ratios are considered (84.4%, 90.0%, and 95.6%). Our proposed two-stage stochastic algorithm adapts the ratio depending on the backlog as shown in Fig. 3.

5. CONCLUDING REMARKS

This paper presents a self-adaptive face identification systems for multi-CCTV based surveillance monitoring applications. The proposed algorithm consists of two stages as follows: For the first stage, the bitrates of all CCTV cameras are adaptively determined based on Lyapunov optimization framework for time-average identification accuracy maximization subject to system queue stability. After time-average optimally computing the bitrates of all connected multi-CCTV cameras, proportional compression ratio scaling is utilized for multi-CCTV camera scheduling and coordination.

6. ACKNOWLEDGEMENT

This research was supported by National Research Foundation of Korea (2016R1C1B1015406). Joongheon Kim is a corresponding author (email: `joongheon@cau.ac.kr`).

7. REFERENCES

[1] J. Dams, M. Hoefer, and T. Kesselheim, "Jamming-Resistant Learning in Wireless Networks," *IEEE/ACM Transactions on Networking*, 24(5):2809–2818, October 2016.

[2] D. Ravi, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Z. Yang, "Deep Learning for Health Informatics," *IEEE Journal of Biomedical and Health Informatics*, 21(1):4–21, Jan. 2017.

[3] O. Arandjelovic, D. S. Pham, and S. Venkatesh, "CCTV Scene Perspective Distortion Estimation From Low-Level Motion Features," *IEEE Transactions on Circuits and Systems for Video Technology*, 26(5):939–949, May 2016.

[4] N. Bof, R. Carli, A. Cenedese, and L. Schenato, "Asynchronous Distributed Camera Network Patrolling Under Unreliable Communication," *IEEE Transactions on Automatic Control*, 62(11):5982–5989, November 2017.

[5] M. Kerr and R. van Schyndel, "Adapting Law Enforcement Frameworks to Address the Ethical Problems of CCTV Product Propagation," *IEEE Security Privacy*, 12(4):14–21, July 2014.

[6] G. Thiel, "Automatic CCTV surveillance-towards the VIRTUAL GUARD," *IEEE Aerospace and Electronic Systems Magazine*, 15(7):3–9, July 2000.

[7] Z.-J. Zha, H. Zhang, M. Wang, H. Luan, and T.-S. Chua, "Detecting Group Activities With Multi-Camera Context," *IEEE Transactions on Circuits and Systems for Video Technology*, 23(5):856–869, May 2013.

[8] C. Ding, J. H. Bappy, J. A. Farrell, and A. K. Roy-Chowdhury, "Opportunistic Image Acquisition of Individual and Group Activities in a Distributed Camera Network," *IEEE Transactions on Circuits and Systems for Video Technology*, 27(3):664–672, March 2017.

[9] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*, Morgan & Claypool, 2010.

[10] J. Kim, G. Caire, and A. F. Molisch, "Quality-Aware Streaming and Scheduling for Device-to-Device Video Delivery," *IEEE/ACM Transactions on Networking*, 24(4):2319–2331, August 2016.

[11] J. Kim, F. Meng, P. Chen, H. E. Egilmez, D. Bethanabhotla, A. F. Molisch, M. J. Neely, G. Caire, and A. Ortega, "Demo: Adaptive Video Streaming for Device-to-Device Mobile Platforms," in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, Miami, FL, USA, September 2013.

[12] J. Koo, J. Yi, J. Kim, M. A. Hoque, and S. Choi, "REQUEST: Seamless Dynamic Adaptive Streaming over HTTP for Multi-Homed Smartphone under Resource Constraints," in *Proc. ACM International Confernece on Multimedia (MM)*, Mountain View, CA, USA, October 2017.

[13] M. J. Neely, "Energy Optimal Control for Time Varying Wireless Networks," *IEEE Transactions on Information Theory*, 52(7):2915–2934, July 2006.

[14] S. Moeller, A. Sridharan, B. Krishnamachari, and O. Gnawali, "Routing without Routes: The Backpressure Collection Protocol," in *Proc. ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)* Stockholm, Sweden, April 2010., (Best Paper Award).

[15] Y. J. Mo, W. Lee, D. Nyang, "Dynamic Security-Level Maximization for Stabilized Parallel Deep Learning Architectures in Surveillance Applications," *IEEE Symposium on Privacy-Aware Computing (PAC)*, Washington DC, USA, August 2017.

[16] D. Kim, J. Kim, and J. Y. Bang, "A Reliable, Self-Adaptive Face Identification Framework via Lyapunov Optimization," in *ACM Symposium on Operating Systems Principles (SOSP) AI Systems Workshop (AISys)*, Shanghai, China, October 2017.