VOLUME 3 ISSUE 1

IEEE ComSoc
IEEE Communications Society

IEEE CSR Newsletter

December 2022

Inside This Issue

Welcome to the CSR Newsletter.

Preface

Welcome to the third edition of the IEEE Technical Committee on Communications Switching and Routing (CSR Newsletter). This is the third issue of the IEEE CSR Newsletter. The purpose of this newsletter is to provide specialized information to the members of the CSR and the communications community at large, the newsletter aims at inviting the authors of successful research projects and experts from the committee members and the larger communications community about CSR-related research activities to share their experience and knowledge by contributing in short news. This issue continues to bring to you the latest news of our technical committee by providing information about our meetings, events activities, and members' news. I would like to thank the colleagues who shared with us their latest news. In this volume, we have selected 3 papers on Transfer Learning, Mutli-Access Networks and ML Estimation for Discrete-Time systems.

I hope you will enjoy this issue and I look forward to furthering good members' news to share.

Scott Fowler
CSR Secretary and Newsletter editor

TC-CSR

The objective of the Technical Committee on Communications Switching and Routing (TC CSR) is to advance the state of the art in theory and applications of information switching and networking by the following:

- To bring together professionals with interests in various aspects of Communications Switching and Routing such as theory and architecture (B-ISDN/ATM, optical switching), teletraffic theory, mobility and call control, signaling protocol, Intelligent Networks, and service features, switching software architecture, management, economics, and applications of switching and routing systems.
- To provide a forum for discussion and exchange of technical matters among members in the committee meetings, special interest group meetings, topical meetings, and workshops.
- To sponsor technical sessions at conferences and stimulate high-quality technical papers for conferences and IEEE publications.
- To organize special issues of IEEE publications.

Committee

Executive committee

Chair
Dr. Lotfi Mhamdi
School of Electronic and
Electrical Engineering
The University of Leeds
Leeds, LS2 9JT, UK
Tel: +44 (0)113 343 6919
Tel: +44 (0)113 343 2000
Email: l.mhamdi@leeds.ac.uk

Vice Chair
Dr. Cheng Li
Electrical and Computer Engineering
Faculty of Engineering and Applied Science
Memorial University
St. John's, NL, A1B 3X5, Canada
Tel: +1 709 864 8972
Fax: +1 709 864 4042
Email: licheng@mun.ca

Secretary
Dr. Scott Fowler
Department of Science and Technology (ITN)
Division of Communications and Transport Systems (KTS)
Linköping University, Campus Norrköping
Bredgatan 34, Norrköping, SE-601 74, Sweden
Tel: +46 11 36 32 98
Email: scott.fowler@liu.se

Committee Meetings

Current Meeting:

- Meeting at GC 2022
  Location: Online Meeting
  7:00am–9:00am EST, Tuesday, November 29, 2022

Previous Meeting:

- Meeting at ICC 2022
  Location: Online Meeting
  7:00am–9:00am EST, Tuesday, May 6, 2022

Next Meeting:

- Meeting TBD, hopefully at the same time as time ICC 2023
  Location TBD

Activities

CSR TC Meeting

Date: Tuesday, November 29, 2022

Time: 7:00 AM – 9: 00 AM, Online: Zoom

The agenda for the meeting is:

(1) Welcome and Introduction
(2) Approval of the agenda
(3) Approval of the meeting minutes at ICC 2022
(4) Report from the Chair
(5) TC Special Interest Groups (SIGs) and Newsletter
(6) IEEE ComSoc Student Competition
(7) TC activities at conferences and workshops
  - HPSR
  - Globecom/ICC
(8) Distinguished Service Award Committee
(9) Other business
(10) Adjourn

HPSR Steering Committee Meeting

Time: Tuesday 22nd November 2022, from 1:00 – 2:00 pm

Location: Online on Zoom

Call for interest in TC-CSR SIGs

Topics of interest to the CSR, including:

- Switching and Routing
- Traffic Engineering
- Network Signalling
- Network Software, Hardware, and Middleware
- Packet Network Applications
- Network Processors
- Network and Function Virtualizations
- Data Center Networking and Cloud Computing
- Mobile Networks
- Energy Efficient and Green Networking
- Information Networks
- Intelligent System Analysis/Analytics
- Optimal Systems
- Optimization of Systems

Members News

- **IEEE ComSoc Student Competition 2022**
  - **1st Place**: **The Owl: An Accessible Immersive Telepresence System for the Future of Human Communication**
    Gonzalez Diego, Universidad Carlos III de Madrid, Spain
  - **2nd Place: Internet of Bodies: Digital Holistic Healthcare**
    Alamoudi Abeer, King Abdullah University of Science and Technology, Saudi Arabia

Announcements

Special Interest Groups (SIG)

- Looking for Special Interest Groups
- Please contact the SIG coordinator is Thiago Abreu, University of Paris-Est, Creteil, France, thiago.wanderley-matos-de-abreu@u-pec.fr

Distinguished Technical Achievement Award: Chair and Subcommittee

- **The new Committee is:**
  - **Prof. Abdelhamid Mellouk (Chair)**
  - Prof. Shiwen Mao, Member
  - Prof. Roberto Rojas-Cessa, Member
  - Prof. Stefano Giordano, Member
  - Dr. Scott Fowler, Member

Upcoming Events

Conference	Location	Date
ICC 2023	Rome, Italy	May 28-Jun. 1, 2023
Globecom 2023	Kuala Lumpur, Malaysia	Dec. 2023
ICC 2024	Denver, USA	Jun. 2024
Globecom 2024	Cape Town, South Africa	Dec. 2024

Optimal IoT Machine Learning Estimation for Discrete-Time Linear System

Scott Fowler
Communications and Transport Systems
Department of Science and Technology, Linköping University, Norrköping, Sweden

I. Introduction

The construction industry is currently undergoing critical adjustments that will improve efficiency, well-being, process improvement, and the introduction of new tools, such as the Internet of Things (IoT). These IoT solutions for construction are having a huge impact on how the construction industry pivots. IoT enables each partner to understand what is happening at each stage of the real-time development process, from planning to positive development, post-development, and how the structure functions during administration. Most construction projects are heavily dependent on concrete as a load-bearing material, for instance in foundation structures. The production of Portland cement[1] used as the primary binder in concrete mixtures comprise around 8% of the world's total CO_2 emissions. The cement and concrete industry, therefore, seeks solutions to reduce the CO_2-emissions. One important solution is to replace part of the Portland cement with SCMs (supplementary cementitious materials) such as fly ash or granulated furnace slag. SCMs have typically a significantly lower carbon footprint compared to cement which reduces the total carbon footprint of the concrete mixture. Due to the lower heat production of Climate-Improved Concrete, these mixtures are also more sensitive to colder weather which is particularly important when pouring concrete in the Nordic countries. Thus, there is a major push to develop new digital smart tools to analyze the data gathered by IoT devices that can assist in making correct decisions using analytics. From the data gathered from the IoT, we need to have accurate results, to predict or estimate a value or class, which is the core objective of a machine learning model is to predict or estimate a value or class, which will be close to an accurate value or class. Kalman filtering is a machine learning approach that can predict or estimate a value or class, which will be closure to an accurate value or class.

II. Problem definition

Kalman filters are used to estimate states based on linear dynamical systems in state space format. The process model defines the evolution of the state from time $t-1$ to time t as [1]–[3]:

$$x_t = Fx_{t-1} + Bu_{t-1} + w_{t-1}, \tag{1}$$

where F is the state transition matrix applied to the previous state vector x_{t-1}, B is the control-input matrix applied to the control vector u_{t-1}, and w_{t-1} is the process noise vector that is assumed to be zero-mean Gaussian with the covariance Q_t, i.e., $w_{t-1} \sim \mathcal{N}(0, Q_t)$.

The process model is paired with the measurement model that describes the relationship between the state and the measurement at the current time t as:

$$\mathbf{z}_t = H\mathbf{x}_t + \mathbf{v}_t, \tag{2}$$

where z_t is the measurement vector, H is the measurement matrix, and v_t is the measurement noise vector that is assumed to be zero-mean Gaussian with the covariance R_t, i.e., $v_t \sim \mathcal{N}(0, R_t)$.

The role of the Kalman filter is to provide an estimate of x_t at time t, given the initial estimate of x_0. The series of measurements, $z_1, z_2, \ldots, z_t$, and the information of the system is described by F, B, H, Q, and R. Note that subscripts to these matrices are omitted here by assuming that they are invariant over time as in most applications. Although the covariance matrices are supposed to reflect the statistics of the noises, the true statistics of the noises are not known or not Gaussian in many practical applications. Therefore, Q and R are usually used as tuning parameters that the user can adjust to get desired performance.

III. Kalman filter algorithm

Kalman Filter is an optimal estimation algorithm that can be measured indirectly to find the best estimate of states by combining measurements from various sensors in the presence of noise. In the Kalman filter, there are two main steps that we need to perform: 1) the Prediction step and 2) the Update step. We perform the prediction step to compute our new belief state after we apply a control signal u_t to the following motion model. After we applied the control signal u_t, (we assume to have received a sensor measurement, z_t) we perform the state based update step received from the sensor measurement. The Kalman filter algorithm is summarized as follows[2]:

[1]Generic term for the type of cement used in virtually all concrete.

[2]Terms "prediction" and "update" are sometimes referred to as "propagation" and "correction," respectively

Prediction step:

Predicted state estimate	$\hat{x}_t^- = F\hat{x}_{t-1}^+ + Bu_{t-1}$ (3)
Predicted error covariance	$P_t^- = FP_{t-1}^+F^T + Q_k$ (4)

Update step:

Measurement residual	$\widetilde{y}_t = z_t - H\hat{x}_t^-$ (5)
Kalman gain	$K_t = P_t^- H^T \left(R + HP_t^- H^T\right)^{-1}$ (6)
Updated state estimate	$\hat{x}_t^+ = \hat{x}_t^- + K_t\widetilde{y}_t$ (7)
Updated error covariance	$P_t^+ = (I - K_t H)\, P_t^-$ (8)

In the above equations, the hat operator, ^, means an estimate of a variable, that is, $\hat{x}$ is an estimate of x. The superscripts $-$ and $+$ denote predicted (prior) and updated (posterior) estimates, respectively. The predicted state estimate evolved from the previous updated state estimate. The new term P is called state error covariance.

In the update stage, the measurement residual $\widetilde{y}_t$ is computed first. The measurement residual, (also known as innovation), is the difference between the true measurement, z_t, and the estimated measurement, $H\hat{x}_t^-$. The filter estimates the current measurement by multiplying the predicted state by the measurement matrix. The residual, $\widetilde{y}_t$, is later then multiplied by the Kalman gain, K_t, to provide the correction, $K_t\widetilde{y}_t$, to the predicted estimate $\hat{x}_t^-$. After it obtains the updated state estimate, the Kalman filter calculates the updated error covariance, P_t^+, which will be used in the next time step. Note that the updated error covariance is smaller than the predicted error covariance, which means the filter is more certain of the state estimate after the measurement is utilized in the update stage.

We need an initialization stage to implement the Kalman filter. As initial values, we need the initial guess of state estimate, $\hat{x}_0^+$, and the initial guess of the error covariance matrix, P_0^+. Together with Q and R, $\hat{x}_0^+$ and P_0^+ play an important role to obtain desired performance. One can obtain implement a Kalman filter by implementing the prediction and update stages for each time step, $t = 1, 2, 3, \ldots$, after the initialization of estimates.

IV. Results

In Figure 1 shows the results when the Kalman Filter is used to estimate the state of the data. This shows how much the Kalman filter is cleaning out the noisy readings (Black). We can also see how one of the sensors (Blue) has more variation. Since the Kalman Filter adjusted for the sensor's noise independently means, the Kalman filter does not just clean up the data measurements but also projects them onto the state estimate. With this feature, the variance is reduced to 2.6, the standard deviation is reduced to 1.6, a percentage reduction to 77.6% and 88.1%. In addition, regarding the missing data during (such as around 1500, 4500, etc,) the Kalman Filter can estimate the status of this missing value, thereby improving the continuity and reliability of the control system.

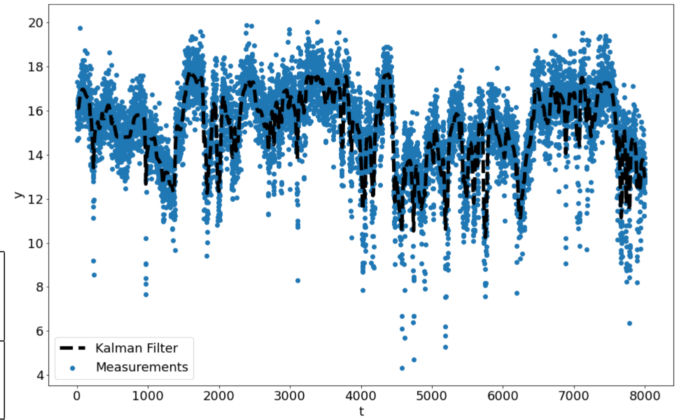

Figure 1: *Prediction results using Kalman Filter.*

V. Conclusion

Kalman Filter algorithm is a great way to model and predict certain types of data. Specifically, the data should have some state-varying or time-varying component. These learning modules take appropriate parameters as input/internal actuator operational status, to predict estimated error in sensor readings as output. A simple ML Kalman Filter analysis of the results indicates that the proposed learning-to-prediction model has significantly improved the accuracy. Further development is the integration between the IoT systems and the working method of Building Information Modeling (BIM). For this purpose, the building model is supplemented with the current IoT data, which allows a timely comparison between the planned and actual work. This can be created in the building model and provided with information about the weather, concrete receipts, environment data, other ML models, and personal protective measures (health and safety) that could be taken into future consideration.

Acknowledgement

The project is funded by the strategic innovation programme Smart Built Environment (Project number: 2021-0030)

References

[1] Youngjoo Kim and Hyochoong Bang. Introduction to kalman filter and its applications. In Felix Govaers, editor, *Introduction and Implementations of the Kalman Filter*, chapter 2. IntechOpen, Rijeka, 2018.

[2] Youngjoo Kim. Introduction to kalman filter and its applications. https://uk.mathworks.com/matlabcentral/fileexchange/68262-introduction-to-kalman-filter-and-its-applications, 2018.

[3] Rey Wiyatno. Kalman filter. https://github.com/rrwiyatn/robotik/tree/main/kalman_filter, 2021.

Transfer learning based urban river quality prediction

Tharsana Balachandran*, Thiago Abreu*, Manel Naloufi*†, Sami Souihi*, Françoise Lucas† and Aurélie Janne‡
*Image, Signal and Intelligent Systems (LiSSi) Laboratory, University of Paris-Est Créteil, France
†Laboratoire Eau, Environnement et Systèmes Urbains (Leesu) Université Paris-Est Créteil, École des Ponts ParisTech, France
‡Syndicat Marne Vive, France

I. Introduction

The smart environment field has a great importance within the new smart city paradigm. It allows the management of natural resources by decision-makers, given data collected through sensor networks and the Internet of Objects (IoT) associated with machine learning approaches. One interesting use case concerns the estimation of surface water quality of urban rivers, a topic currently approached by the towns within the basin of the Marne River in the suburbs of Paris, France.

Currently, the water quality forbids the bathing in the Marne, but local associations and public administrations intend to open future bathing sites by the river. Therefore, their goal is to restore good water quality in the river, and to ensure their maintenance by a regular monitoring of physical-chemical characteristics and microbiological pathogens in water, as demanded by the European directives of surface waters (91/271/EEC directive for the treatment of urban waste water, 2000/60/EC and 2006/7/EC directives for the management of bathing water quality). These directives take into account the presence of two groups of non-pathogenic fecal indicator bacteria (FIB), *Escherichia coli* (*E. coli*) and intestinal enterococci (except for some strains) in water. Their presence may indicate other enteric pathogens of fecal origin (domestic sewage discharge, animal dejection, runoff loaded with sewage or runoff from pastures), which could increase the risks of gastroenteritis in humans [1].

One current major problem, however, is that monitoring water quality relies on samples that require further analysis in specialized laboratories. This procedure takes time (up to days) and may be very costly for small towns, which may impact the correct assessment of the quality and prevents a real-time follow-up by decision-makers.

Thus, the creation of new tools for real time and cheaper surface water monitoring is an ongoing issue. One solution consists in the use of machine learning approaches, based on physical-chemical data obtained through IoT networks, for the prediction of FIB concentrations in surface water.

II. Machine learning for water quality estimation

Machine learning (ML) methods for water quality monitoring and prediction have been used with good performance when compared to traditional statistical and deterministic models. [2]. These models can be used to predict FIB concentration in surface water, potential bathing site as ”bathable” or ”non-bathable” according to the European directives.

Several problems, however, arise from the use of ML. To create a model, a lot of resources and computation time may be required. Moreover, an important volume of data is often necessary to obtain acceptable results, and the available datasets may not be enough. In the case of Marne, the available data for this research is, unfortunately, still limited, although our research group has been taking several approaches to solve this issue (increasing the number of sampling stations and using IoT devices and sensors to collect data). Therefore, one solution consists in pre-training ML models using datasets from different rivers, which have similar physical-chemical compositions as the Marne. The acquired knowledge is, hence, reused in our target use case to better predict FIB concentrations. This approach is a branch of the ML field, known as transfer learning.

The available dataset from Marne comes from sampling campaigns held every summer since 2015 (except 2016) in order to assess the quality of the Marne River, using 17 monitoring stations alongside the river. In total, we have 1696 samples at the level of the Marne for 5 years (2015, 2017-2020) [3]. Given the available amount of data, we fixed as our goal to create predictive models of only *E. coli* (MPN/100ml) concentrations, since in the Marne River *E. coli* is a more degrading parameter than intestinal enterococci considering the guidelines of the European bathing directives [3].

The final ML model is trained by using a set of physical-chemical parameters from other rivers containing larger datasets, and further enhanced with the data from the Marne. At first, we tried to access French databases to use data of local rivers to pre-train different ML models that we test (to be discussed below), but either not enough data was available, or the parameters were different. After some research, datasets from the environmental platforms of New Zealand (NZ) provided enough data from rivers with similar characteristics and similar data distribution as the Marne (totaling 5,450 samples). In our case, we use temperature (○C), conductivity (µS/cm), turbidity (NFU), suspended solids (SS) (mg/L), ammonium (NH_4^+) (mg-N/L) and total Kjeldahl nitrogen (NTK) (mg-N/L) as the input variables of our ML models, since those characteristics are common to both datasets.

After performing data cleaning of all datasets (incomplete

Marne	DT	KNN	SVM	RF	Boost(DT)	Bag(RF)	VR
RMSE	0.90	0.77	0.81	0.76	0.86	0.74	0.78
MAE	0.37	0.28	0.27	0.29	0.30	0.29	0.36
RPD	1.04	1.21	1.15	1.22	1.23	1.25	1.19

TABLE I

EVALUATION OF CREATED MODELS USING DATA FROM THE MARNE

Model	Model 1	Model 2
RMSE	0.61	0.62
MAE	0.33	0.25
RPD	1.38	1.47

TABLE III

EVALUATION OF MODELS 1 AND 2 PERFORMANCE WITH MARNE DATA

data, outliers, highly correlated features and redundancy), we separate the samples randomly (80% for training, 20% for testing), with a standardization of the parameters.

Given that our data is labeled, we focused on supervised learning approaches to classify bathing sites as "bathable" or "non-bathable" in terms of *E. coli* concentration values. We choose to test i) traditional algorithms, such as Decision Trees (DT), Support Vector Machine (SVM) and K-nearest neighbors (KNN); ii) and also the so-called ensemble algorithms, such as Random Forest (RF), Adaboost and Bagging. Furthermore, we average the predictions of the ML models using voting regression (VR) to ensure better precision. For further details on the algorithms, please refer to [4].

Finally, to evaluate the performance of the different models, we choose three metrics: the root mean square error (RMSE), mean absolute error (MAE) and the ratio of performance to deviation (RPD) [5]. With the RMSE, we can evaluate the standard deviation of the distribution of prediction errors around the line of good prediction. The MAE gives the magnitude of the average error between the predicted values and the actual observations. The RPD is a metric where values < 1.4 indicate that unreliable models; values between 1.4 and 2 indicate moderately accurate models and values higher than 2 indicates good predictive abilities.

III. RESULTS

If we train our ML models using only the data from the Marne River to predict *E. coli* concentration values, we find the values shown in Table I. The RPD analysis shows that none are reliable (RPD ≤ 1.4). A similar analysis is conducted using only the data from New Zealand, with similar results (RPD ≤ 1.4 for all models) (see Table II). In both cases, we can see that the VR models are also under performing. Thus, we decide to create a new Model 1 using the average of both VR models, to ensure the share of knowledge among models using different datasets. Furthermore, we create a Model 2 that takes into account only the five best performing ML models in terms of low RMSE. A more reliable solution is obtained, as shown in Table III, with Model 2 presenting RPD ≥ 1.4.

NZ	DT	KNN	SVM	RF	Boost(DT)	Bag(RF)	VR
RMSE	0.12	0.12	0.14	0.22	0.66	0.24	0.27
MAE	0.05	0.04	0.09	0.05	0.06	0.05	0.06
RPD	0.99	0.97	0.85	0.48	0.17	0.48	0.46

TABLE II

EVALUATION OF CREATED MODELS USING DATA FROM NEW ZEALAND

IV. CONCLUSIONS

Opening bathing sites in urban rivers is a project from different cities, in order to increase life quality of their citizens. This represents, however, an important challenge in terms of improving water quality to reduce the risks of contamination of bathers due to enteric pathogens. Therefore, a continuous monitoring of bacteriological levels is mandatory for such bathing sites, but traditional techniques are not suitable. They often rely on expensive tools or slow sample evaluations, where results are only know days after contamination episodes.

We show in this work that using machine learning techniques and transfer learning can enhance the accuracy of predictive models. We use data from New Zealand's rivers alongside with the Marne ones to create a more accurate model for the French river. Although the results are encouraging, the lack of samples remains problematic.

Therefore, future advances must consider the development of sensing networks and the use of the Internet of Things to increase datasets. Works on the field are currently held by our research team, by mixing the existing high-accuracy devices (which are scarce) and low-cost devices (lower quality, but more data), using the Knowflow platform [6]. Moreover, researchers must investigate sensors deployment techniques, alongside with transmission technologies, such as the use of Low-Power Wide Area Networks (LPWAN) like LoraWAN, Sigfox, NB-IoT, or even 5G. [7].

REFERENCES

[1] J. A. Soller, M. E. Schoen, T. Bartrand, J. E. Ravenscroft, and N. J. Ashbolt, "Estimated human health risks from exposure to recreational waters impacted by human and non-human sources of faecal contamination," *Water research*, vol. 44, no. 16, pp. 4674–4691, 2010.

[2] K. Chen, H. Chen, C. Zhou, Y. Huang, X. Qi, R. Shen, F. Liu, M. Zuo, X. Zou, J. Wang, Y. Zhang, D. Chen, X. Chen, Y. Deng, and H. Ren, "Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data," *Water research (Oxford)*, vol. 171, pp. 115 454–115 454, 2020.

[3] F. Lucas, P. Servais, and A. Janne, "Qualité bactériologique de la zone aval de la Marne : Synthèse des campagnes estivales 2015-2019." *Rapport (in French)*, 2020.

[4] A. Geron, *Hands-on machine learning with Scikit-Learn, Keras and TensorFlow : concepts, tools, and techniques to build intelligent systems*, 2nd ed. Sebastopol (Calif.): O'Reilly Media, 2019.

[5] M. Naloufi, F. S. Lucas, S. Souihi, P. Servais, A. Janne, and T. Abreu, "Evaluating the performance of machine learning approaches to predict the microbial quality of surface waters and to optimize the sampling effort," *Water*, vol. 13, no. 18, 2021.

[6] "Public Lab KnowFLow," https://publiclab.org/wiki/knowflow, accessed on 15 November 2022.

[7] H. Rahimi, A. Zibaeenejad, and A. A. Safavi, "A novel IoT architecture based on 5G-IoT and next generation technologies," in *In Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),Vancouver, Canada*, 2018, pp. 81–88.

Multi-Access Networks Management in Dynamic Heterogeneous Environment

Tran-Tuan CHU*, Mohamed Aymen LABIOD*, Brice AUGUSTIN*, and Abdelhamid MELLOUK*
* University of Paris-Est Creteil, LISSI, TincNET (CIR), F-94400, Vitry-sur-Seine, France
Email: tuan.chu-tran@u-pec.fr, mohamed-aymen.labiod@u-pec.fr, brice.augustin@u-pec.fr, mellouk@u-pec.fr

I. Introduction

TCP and UDP have been the only transport protocols in the pervasive TCP/IP stack for decades. However, they were designed with the underlying presumption that only one path could be active between two end hosts at a time. We can see how obsolete this idea has become just by noting the abundance of network interfaces that today's end hosts are equipped with. To allow the simultaneous use of different paths that may be present between a transmitter and a receiver, multipath transport protocols have been developed. Any such protocol can bundle different network interfaces and switch between them transparently. The most popular implementations to date are Multipath TCP (MPTCP), Concurrent Multipath Transfer for SCTP (CMT-SCTP), and Multipath QUIC (MPQUIC). These transport protocols enable the end-user device to have better throughput, lower latency, and a fallback solution in network failure situations by simultaneously exploiting several accessible network interfaces.

On the other hand, a plethora of technologies currently coexists in the networking landscape. For instance, to realize the Internet of Everything, 5G aims to integrate diverse technologies, such as vehicular networking, machine-to-machine (M2M) communications, device-to-device (D2D) communications, non-terrestrial networks (NTNs), mobile edge computing (MEC), Internet-of-Things (IoT), cloud radio access networks (CRANs), cloud computing, unmanned aerial vehicles (UAVs), etc. From the perspective of network architecture, wireless networks evolved from homogeneous networks (HomNets) to heterogeneous networks (HetNets). In many circumstances, as in the examples listed, multiple interfaces are available to end users, providing the ability to use multiple available network interfaces simultaneously. Thereby, in order to make the most of the multipath approach, research focuses on three orthogonal aspects: congestion control, path management, and scheduling. Indeed, the multipath scheduler plays a key role as it is responsible for distributing data over multiple paths and therefore affects the overall network performance. Poor scheduling decisions can increase out-of-order packet arrival or introduce Head-of-Line Blocking (HoLB). The 3rd Generation Partnership Project (3GPP) has also emphasized this potential cooperation in the Technical Specification (TS) 23.501 (Release 16) by claiming the Access Traffic Steering, Switching, and Splitting (ATSSS) architecture could be enhanced by multipath transport protocols [1].

II. Multipath scheduling performance over heterogeneous paths

In order to avoid congestion in the multipath environment, a multipath scheduling algorithm selects which path should be used to send the packet to make the best use of the available paths. Multi-path schedules can be divided into two categories: those based on predetermined rules and those based on Machine Learning (ML).

One of the most basic schedulers is the Round-robin (RR) scheduling algorithm, which sends data over each available path sequentially as long as there is space in the congestion window (CWND). The second basic scheduler is the minRTT scheduling algorithm, which is the default algorithm in MPTCP [2] and MPQUIC [3] and it is based solely on the round trip time (RTT). Blocking Estimation Scheduler (BLEST) [4] and Earliest Completion First (ECF) [5] just like the first two schedulers are also based on predefined rules. They propose, with different approaches, a waiting mechanism to delay the emission if only the path with the highest RTT is available. Several other schedulers based on predefined rules exist but will not be mentioned here.

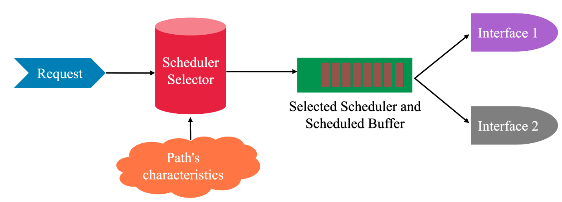

Fig. 1. GaDAM proposed Architecture

On the other hand, we observe an increasing interest in the design of ML-based multipath schedulers to overcome the performance limitations of rule-based schedulers. Peekaboo [6], which is a learning-based scheduler, applies a stochastic adjustment method for the MPQUIC protocol after choosing a deterministic strategy based on the Linear Upper Confidence Bound (LinUCB). GADaM offers another approach by positioning itself as a scheduler selector [7]. After an analysis that showed that in the same scenario with varying network conditions, none of the schedulers mentioned above could systematically present the best performance. Therefore, the authors proposed a deep-learning-based approach that selects the most adapted scheduler, given the network conditions observed in a previous time window. Fig.1 indicates the

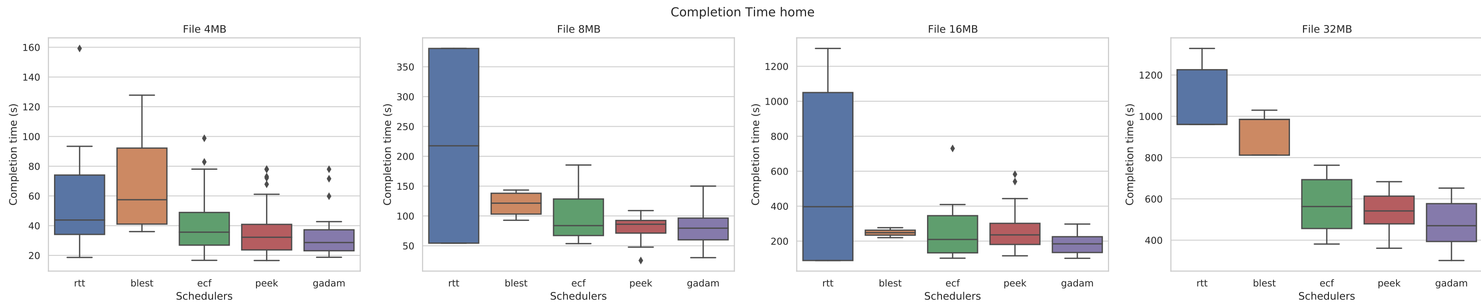

Fig. 2. Schedulers Performances Comparison

suggested architecture to implement the proposal paradigm "scheduler selector" in the practical network stack. Otherwise, when it comes to performance evaluation, due to the time required, real-world experimental evaluations are more the exception than the rule, and most validations are done in simulated environments.

III. Performance evaluation

In order to fully understand the actual performance of each scheduler in HetNets, experiments were conducted on a set of bare-metal equipment. The general architecture is composed of three main elements: a cluster of servers, the network infrastructure, and a cluster of clients. These have multiple interfaces and have been used to determine the actual performance of each scheduler in HetNets. A complete description of these environments can be found in [8]. As shown in 2 while minRTT seemed to have the slowest download completion time with the highest variance, BLEST and ECF retained their rankings. Whereas, Peekaboo performed admirably when compared to its predecessors. However, GADaM has the best average download completion time and a noticeably reduced variance when compared to the most cutting-edge Peekaboo in 4MB, 16MB and 32MB cases. GADaM's approach could be considered to outperform the predefined state-of-the-art scheduler suggestion.

IV. Conclusion

The emergence of various technologies and applications makes networks increasingly heterogeneous. Aside from the ongoing solutions for better frequency spectrum utilization, particularly through Software Defined Radio (SDR) solutions, the use of multiple network access is a cutting-edge solution. This is all the more true with multipath transport protocols, provided that scheduling is correctly dealt with at these protocols' level.

Regardless, the deep learning-based schedulers' advantage has been proven against the deterministic approaches in both simulated and practical experiments. Indeed, GaDAM which is a deep-learning-based multipath scheduler selector, outperforms the rest of the approaches. The main idea of GADaM is to select the most suitable scheduler, which means that GADaM performance strongly depends on the performance of its candidate. These remarks also pointed out that the GADaM improvement could be much more significant under frequently changing network conditions. However, further investigation and in-depth analysis should be considered for future work to better understand.

Acknowledgment

This research is funded by MESRI France (French Higher Educational, Research and Innovation Ministry) under the doctoral grant (Allocation de these - UPEC). Otherwise, the authors would like to thank Mr. Papa Bara DIAKHATE for his contribution to the data acquisition.

References

[1] 3GPP. 23.501: System architecture for the 5g system v16.4, 03.2020.

[2] Alan Ford, Costin Raiciu, Mark Handley, and Olivier Bonaventure. Tcp extensions for multipath operation with multiple addresses. *RFC*, 8684:1–68, 2013.

[3] Quentin De Coninck and Olivier Bonaventure. Multipath QUIC: Design and evaluation. *CoNEXT 2017 - Proceedings of the 2017 13th International Conference on emerging Networking EXperiments and Technologies*, pages 160–166, 2017.

[4] Simone Ferlin, Ozgu Alay, Olivier Mehani, and Roksana Boreli. BLEST: Blocking estimation-based MPTCP scheduler for heterogeneous networks. *2016 IFIP Networking Conference (IFIP Networking) and Workshops, IFIP Networking 2016*, pages 431–439, 2016.

[5] Yeon Sup Lim, Erich M. Nahum, Don Towsley, and Richard J. Gibbens. ECF: An MPTCP Path Scheduler to Manage Heterogeneous Paths. *Performance Evaluation Review*, 45(1):33–34, 2017.

[6] Hongjia Wu, Ozgu Alay, Anna Brunstrom, Simone Ferlin, and Giuseppe Caso. Peekaboo: Learning-Based Multipath Scheduling for Dynamic Heterogeneous Environments. *IEEE Journal on Selected Areas in Communications*, 38(10):2295–2310, 2020.

[7] Tran-Tuan Chu, Mohamed Aymen Labiod, Hai-Anh Tran, and Abdelhamid Mellouk. Gadam: Generic adaptive deep-learning-based multipath scheduler selector for dynamic heterogeneous environment. In *ICC 2022-IEEE International Conference on Communications*, pages 4908–4913. IEEE, 2022.

[8] Papa Bara Diakhate, Tran-Tuan Chu, Mohamed Aymen Labiod, Brice Augustin, Hai-Anh Tran, and Abdelhamid Mellouk. Experimental evaluation of multiple multipath schedulers over various urban mobile environments. In *Proceedings of the eleventh International Symposium on Information and Communication Technology*, 2022.