

Learning Wi-Fi Connection Loss Predictions for Seamless Vertical Handovers Using Multipath TCP

Jonas Höchst^{*†}, Artur Sterz^{*†}, Alexander Frömmgen^{*}, Denny Stohr^{*}, Ralf Steinmetz^{*}, Bernd Freisleben^{*†}

^{*}Dept. of Electrical Engineering & Information Technology, TU Darmstadt, Germany

E-mail: {jonas.hoechst, artur.sterz, alexander.froemmgen, denny.stohr, ralf.steinmetz, bernd.freisleben}@maki.tu-darmstadt.de

[†]Dept. of Mathematics & Computer Science, Philipps-Universität Marburg, Germany

E-mail: {hoechst, sterz, freisleb}@informatik.uni-marburg.de

Abstract—We present a novel data-driven approach to perform smooth Wi-Fi/cellular handovers on smartphones. Our approach relies on data provided by multiple smartphone sensors (e.g., Wi-Fi RSSI, acceleration, compass, step counter, air pressure) to predict Wi-Fi connection loss and uses Multipath TCP to dynamically switch between different connectivity modes. We train a random forest classifier and an artificial neural network on real-world sensor data collected by five smartphone users over a period of three months. The trained models are executed on smartphones to reliably predict Wi-Fi connection loss 15 seconds ahead of time, with a precision of up to 0.97 and a recall of up to 0.98. Furthermore, we present results for four DASH video streaming experiments that run on a Nexus 5 smartphone using available Wi-Fi/cellular networks. The neural network predictions for Wi-Fi connection loss are used to establish MPTCP subflows on the cellular link. The experiments show that our approach provides seamless wireless connectivity, improves quality of experience of DASH video streaming, and requires less cellular data compared to handover approaches without Wi-Fi connection loss predictions.

I. INTRODUCTION

Smartphones have become our daily mobile companions to provide wireless access to communication, information, and entertainment services. Since a large amount of data is not downloaded in advance but streamed on demand via the Internet, seamless connectivity using both Wi-Fi and cellular interfaces is desirable. The mobility of smartphone users leads to the problem of deciding when to use which wireless connection. Most smartphones use Wi-Fi as their default interface, since many cellular data plans will be throttled after exceeding a certain limit. The decision when to perform vertical handovers is often based on the Wi-Fi received signal strength indicator (RSSI) and timeouts for the transmission of packets. However, if a user is leaving a location while listening to a music stream or watching a streamed video, the established Wi-Fi connection eventually becomes unavailable and the streaming process stops. The mobile operating system detects the connection loss some time after the connection is lost. Finally, the application needs to reestablish the connection.

In this paper, we present a novel approach to predict Wi-Fi connection loss before the connection breaks to perform seamless vertical Wi-Fi/cellular handovers. Our approach re-

lies on data collected by multiple smartphone sensors (e.g., Wi-Fi RSSI, acceleration, compass, step counter, air pressure) to predict Wi-Fi connection loss and uses Multipath TCP (MPTCP) to dynamically switch between different wireless connectivity modes. We train a random forest classifier and an artificial neural network on roughly 20 GB of sensor data collected by five smartphone users over a period of three months. The trained models are efficiently executed on smartphones and reliably predict Wi-Fi connection loss 15 seconds ahead of time, with a precision of up to 0.97 and a recall of up to 0.98. Furthermore, we present results of four DASH video streaming experiments that run on an Android smartphone and make use of available Wi-Fi/cellular networks. The neural network predictions for Wi-Fi connection loss are used to establish MPTCP subflows on the cellular link. Our experiments show that the proposed approach provides seamless wireless connectivity, improves quality of experience by increasing mean opinion scores (MOS) from 2.7 to up to 3.8 for certain scenarios, and requires up to 50% less cellular data compared to handover approaches without Wi-Fi connection loss predictions. The data set, analysis scripts, experimental logs, and the mobile app developed in this paper are publicly available.¹ To summarize, we present:

- a novel approach to predict Wi-Fi connection loss for performing seamless vertical handovers,
- a neural network to learn and predict Wi-Fi connection loss based on a novel combination of smartphone sensors,
- a vertical handover method that uses MPTCP for switching between wireless connection modes at runtime,
- an implementation on off-the-shelf smartphones to demonstrate its performance in real-world scenarios; our results show significant improvements in terms of Quality of Experience and the amount of cellular data consumed.

The paper is organized as follows. Section II discusses related work. Section III presents an overview of our approach. Section IV analyzes the methods for performing predictions, and Section V evaluates the performance using vertical Wi-Fi cellular handovers on smartphones. Section VI concludes the paper and outlines areas for future work.

¹<https://umr-ds.github.io/seamcon/>

II. RELATED WORK

A. Predicting Wi-Fi Connection Loss

Several approaches to predict Wi-Fi connection loss for performing handovers have been proposed in the literature [1]. Nasser et al. [13] use neural networks to predict Wi-Fi connection loss events based on RSSI. Horich et al. [6] use a fuzzy logic controller (FLC) for making decisions about performing handovers, where the parameters for the FLC are learned using a neural network.

Lin et al. [9] propose to use standard Wi-Fi connection properties and a neural network to predict Wi-Fi connection loss. Monsour et al. [11] use a combination of user velocity and the Allan variance of the RSSI to predict Wi-Fi connection loss, and use PMIPv6 to manage the predicted Wi-Fi connection loss. Khan et al. [7] propose a fuzzy logic system to predict Wi-Fi loss events based on various parameters, such as delay, jitter, bit error rate, packet loss, communication cost, response time, and network load.

These approaches are limited to information of wireless connections, which may be helpful to create metrics for Wi-Fi quality, but is not always the best information for predicting Wi-Fi connection loss. In contrast, our approach considers information from a wide range of smartphone sensors that indicate the usage context, leading to high-quality predictions.

Other approaches incorporate the mobility of the users [14] or higher level features like social group affiliation, time-of-day, and average duration a user spends in a particular network [21].

The predictions in all of these approaches depend on external factors and indicators. In contrast, our approach only requires information that every current mobile device provides and thus can be used in a straightforward, economically attractive manner. To best of our knowledge, there is no work that uses smartphone sensor data to predict Wi-Fi connection loss.

B. Performing Vertical Handovers

There are extensions to the traditional Internet Protocol that allow users to keep a session alive when (vertical) handovers are performed [17]. These approaches are based on home and foreign agents that forward traffic for the mobile host. Although they are around for a long time, mobile IP is not supported widely. Ma et al. [10] propose a vertical handover method based on the Stream Control Transmission Protocol. While the proposed method is network-independent and thus does not require home and foreign agents, traditional TCP-based applications cannot benefit from the advancements. MPTCP is a TCP extension supporting multiple subflows for a single TCP connection [5]. MPTCP improves throughput and reliability in data center and mobile environments [19], [2]. Paasch et al. [15] evaluate MPTCP as a vertical handover mechanism. The authors propose three MPTCP modes for handover scenarios, namely *Full*, *Backup*, and *Single-Path Mode*. The first two modes maintain subflows on all interfaces, while the *Single-Path Mode* exploits the break-before-make

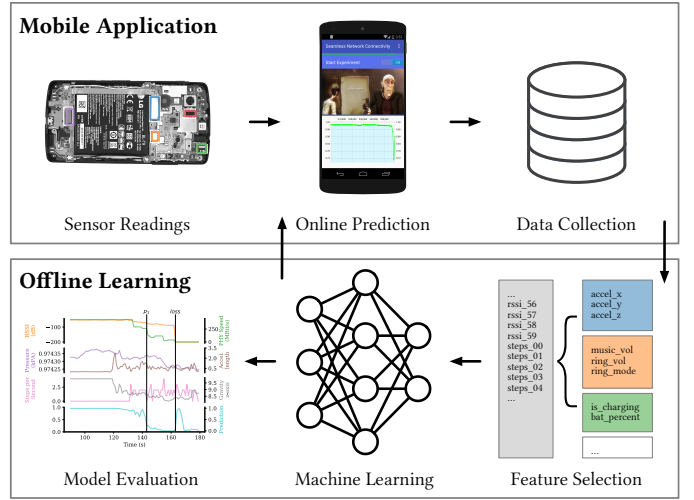


Figure 1: Mobile application and offline learning.

design of MPTCP. Pluntke et al. [18] use MPTCP as a vertical handover mechanism to shift connections between cellular and WiFi connectivity and finally to save energy. De Coninck and Bonaventure [4] further improve the handover by speeding up packet retransmissions after the cellular subflow is established.

The handover mechanisms in these approaches are either reactive, resulting in temporary connection losses, or use redundancy, leading to high bandwidth consumption, which is often contrary to the users' preferences.

III. CONCEPTUAL OVERVIEW

Figure 1 shows the components of our approach and the workflow. First, raw sensor data is collected by a mobile app developed for our work and uploaded to a server for further processing. The raw data is appropriately preprocessed and enriched with additional higher level features. The resulting data is then used to train and evaluate a random forest classifier and different neural network architectures. The data preprocessing operations as well as the trained models are transpiled to Java code and integrated into the mobile app on the smartphone, which in turn makes online predictions for Wi-Fi connection loss 15 seconds ahead of time. Based on these predictions, vertical Wi-Fi/cellular handover is performed using MPTCP. We explain the main steps of our approach in more detail below. Neural network model building and Wi-Fi connection loss predictions for performing MPTCP handovers on a smartphone are discussed in Sections IV and V, respectively.

A. Smartphone Sensors

Modern smartphones offer a variety of sensors that directly or indirectly measure different properties, as explained below. Even though every individual feature might not be a good Wi-Fi connection loss indicator, combinations of seemingly irrelevant features can improve the prediction accuracy.

Motion: Depending on the abstraction level, direct motion sensor readings (e.g., accelerometer, gyroscope, magnetometer), sensor readings cleaned from unwanted influences (e.g., gravity, linear acceleration, rotation vector), or higher level sensor readings as hardware processed triggers (e.g., significant motion, step counter, step detection), are good predictors for user movement.

Orientation: Orientation sensors can reveal more specific situations, where a phone is in the pocket or laying on a table. The proximity sensor is typically used to detect whether the smartphone is held to the ear, but can also be a good hint for other situations, e.g., to detect whether the smartphone is face down on the table.

Environment: Environmental sensors include sensors for measuring ambient light to control screen brightness, humidity, air pressure, and ambient temperature. Rapid changes in these sensor readings can reveal a sudden change of the smartphone situation, e.g., going outdoors.

Global position: GPS can be useful in combination with a world map of Wi-Fi availability. Due to quality concerns with indoor GPS traces and high energy consumption, we discarded GPS in our work.

User interactivity: The user's current context can be derived from various indicators, like device state (interactive, idle, power save), current charging state, audio state (speaker, headphones, and their volumes), and ringer mode.

Wi-Fi properties: Wi-Fi properties, obtained by the radio interface, provide insights into the current connection quality along with reachable networks. Relevant indicators include RSSI, data link layer speed, and used frequency bands.

B. Sensor Data Preprocessing

To learn Wi-Fi connection loss predictions, the sensor data needs to be preprocessed. The time component of the sensor readings needs to be incorporated in the feature vector.

Sensor sampling: The used heterogeneous sensors have different reading frequencies. Motion and orientation sensors can be read with a rather high sampling rate R of 50 Hz, while other sensors are available and useful just under 1 Hz. As a trade-off between energy consumption and sensor data quality, we chose a sensor data sampling rate of $R = 1$ sample per second. Sensors with lower sampling rates are filled until a new value becomes available.

Observation and prediction window: To enrich the discrete sensor readings and to consider the temporal component, the sensor readings are processed in an observation window OW . We use an observation window of 60 seconds, which is derived from common walking speeds and Wi-Fi access point ranges. The earlier a Wi-Fi connection loss is predicted, the more effective the transition between Wi-Fi/cellular is. To define an upper bound on the prediction window, the quality characteristics of the used network protocols are important. Transport protocols, such as TCP, use slow-start to avoid congestion. To compensate for this low-bandwidth start, an early prediction is useful. As a trade-off between performance and farsightedness, and to avoid long-running redundant MPTCP

connections, which are energy and data plan consuming, we use a prediction window of up to 15 seconds.

Feature vector: The feature vector presented to the learning algorithm consists of the sensor readings in the observation window. Each individual sensor contributes $OW \times SR$ values to the feature vector. In the selected configuration, this results in 60 values per sensor. Furthermore, all features are normalized by removing the mean and scaling to unit variance, as required for the machine learning algorithms used in our approach.

C. Precision and Recall

When the Wi-Fi connection loss is predicted too early or too often, this can result in higher consumption of the commonly restricted data plans of the users. Predicting it too late, on the other hand, can result in a dissatisfactory QoE. The primary goal is to reach a high recall in predicting Wi-Fi connection loss. In terms of energy efficiency, the secondary goal is to reach a high precision predicting Wi-Fi connection loss, thus not performing unnecessary handovers.

IV. LEARNING WI-FI LOSS PREDICTIONS

In this section, our novel data-driven approach to predict Wi-Fi connection loss is presented.

A. Data Set

We collected about 20 GB of smartphone sensor data from 5 users, with more than 900,000 unique samples, over a period of three months. The users were advised to let the mobile application run throughout the day, thus the traces contain data from the users' daily lives.

Training and test set: Machine learning methods require separate data sets for training and testing to verify the generalization abilities of the trained models. We investigated different ways of building training and test sets: (a) we randomly split the available samples into, e.g., 70% training and 30% test data, and (b) we split by users, to learn and test with different users.

B. Feature Vectors

All features collected on the smartphones can be used as predictors for Wi-Fi connection loss. We used two feature vector sets, namely the *Full* and the *Reduced Feature Vector*.

Full Feature Vector: The data collected by the different users shows that some features are not available on all devices. The 25 features selected for the full feature vector consist of values of all available sensors: Atmospheric pressure: x , δ ; Linear acceleration: x , y , z , $length$; Step counter: δ ; Power: $is\ charging$, $battery\ percentage$; Gravity: x , y , z ; Gyroscope: $length$; Magnetic field: x , y , z ; Orientation: x , y , z ; Rotation: x , y , z ; Wi-Fi: $frequency$, $speed$, $RSSI$. Thus, the feature vector consists of $25 \times 60 = 1500$ features (i.e., with a 60 seconds observation window).

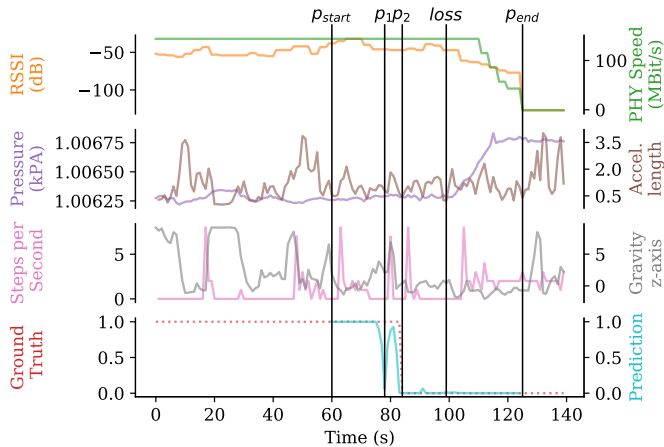


Figure 2: Different sensors leading to an early (p_1) and an ideal (p_2) prediction of Wi-Fi connection loss, based on a trained model with randomly split data.

Reduced Feature Vector: Many of the sensors, like linear acceleration and gyroscope, described in Section III-A share underlying features due to their physical properties. The number of sensors can be reduced by leaving aside these sensors. For the *Reduced Feature Vector*, we used the following sensors: Atmospheric pressure: *delta*; Linear acceleration: *length*; Step counter: *delta*; Power: *is charging*; Gravity: *z*; Wi-Fi: *frequency, speed, RSSI*.

C. Sensor Data Example

Figure 2 shows an example of several sensor data values collected by a smartphone. The figure shows the computed ground truth and a prediction probability value of a neural network based on the *Full Feature Vector*, i.e., a probability value $< 50\%$ means that a Wi-Fi connection loss is predicted and vice versa. The graphical representation of the sensor values shows that no obvious correlation between one of the sensors and the prediction ground truth exists. Nevertheless, each of the sensors shows some information that could be useful. For example, the atmospheric pressure sensor rises from $t = 100$ to $t = 115$, which could be caused by changing the floor in order to leave the building or by a changing ventilation. In combination with the step counter delta, the first option is more likely, also resulting in a higher likelihood for a Wi-Fi connection loss. Another example is the gravity sensor's z axis that reports about 9.81 for the time period from $t = 20$ until $t = 35$, which together with the linear acceleration sensor is a good sign for laying flat on a table. This again reduces the likelihood of a Wi-Fi connection loss event.

For the neural network shown on the bottom in Figure 2, a 60 seconds observation window has to be filled before the first prediction is performed at p_{start} . The classification ends at p_{end} , since the operating system reports that Wi-Fi is unavailable. Since Wi-Fi becomes unavailable at $loss$, the ground truth is 0 from p_2 ongoing, matching the 15 seconds prediction window. The neural network classifier matches the

Table I: *Reduced Feature Vector*, randomly split data, different learners and configurations.

Metric	Forest	NN 1	NN 2	NN 3
Loss Prec.	0.89	0.95	0.97	0.97
Loss Recall	0.98	0.94	0.95	0.95
F_1 -score	0.93	0.94	0.96	0.96

ground truth quite well, with the exception of p_1 , where the classifier predicts the loss slightly too early. This example shows that the combination of sensors available on today's smartphones can lead to an effective prediction of Wi-Fi connection loss.

D. Machine Learning Results

This section presents results of training different methods with the data to predict Wi-Fi connection loss: (a) a random forest classifier [8], and (b) a multi-layer neural network. In particular, we use the MLPClassifier and RandomForest implementations of scikit-learn [16].

Random forest: Since random forest learning depends on equally distributed samples, the data is down-sampled accordingly to match this criterion. The random forest consists of 10 random trees, learned using the Gini criterion. The overall performance of the random forest is satisfactory, since all values are greater than 0.97. However, the precision of the Wi-Fi connection loss class was not very high (0.86), ultimately resulting in triggering early or unnecessary handovers.

RSSI-only neural network: Another basic learning approach is to limit the learner to only use the timeseries of RSSI values, as presented in Section II. During our experiments, different configurations of the neural network were evaluated. The overall performance is comparable to the performance presented in the related work. The classification quality of the Wi-Fi connection loss class did not exceed an F_1 -score of 0.95.

Random Data Split: The results for neural networks learned with randomly split data depend on the neural network architectures. Table I provides an overview of different classifier approaches with the *Reduced Feature Vector*. Classifier *NN 1* consists of 100 hidden neurons, *NN 2* of (300, 200, 100) neurons, and *NN 3* of 5 hidden layers containing (400, 400, 400, 400, 400) neurons. All results were achieved using 70% of the data set exclusively for learning and the remaining 30% for testing. In our experiments, *NN 1* can reach a classification quality comparable to the random forest classifier. The F_1 -score of the Wi-Fi connection loss class reaches up to 0.94, with either a high precision or a high recall, but never both. In general, the negative class, representing stable Wi-Fi connections, is predicted well by all tested neural network classifiers. The experiments show that neural networks can reach both high precision and high recall in the positive Wi-Fi connection loss class.

The results presented in Table I show that *NN 2* and *NN 3* provide reasonably good performance for both precision and

recall in the Wi-Fi connection loss class. Even the neural network *NN 2* consisting of three layers shows significant improvements compared to the flat neural network discussed in the previous paragraph. It reaches an F_1 -score of 0.96 with slightly lower recall or precision.

Other neural network architectures with up to 10 hidden layers were tested. Both precision and recall could not be improved. Splitting the data randomly, *NN 2* and *NN 3* perform equally well and enable a prediction with 97% precision, 95% recall, and a combined F_1 -score of 0.96.

User-based Data Split: When testing for previously unseen users, the precision of the loss worsens in our prediction. With 0.93, 0.92, and 0.79 precision in the Wi-Fi loss class, the *Reduced Feature Vector* generalizes better compared to the *Full Feature Vector* resulting in 0.91, 0.72, and 0.68 precision.

The results show that the neural networks are capable of generalizing even among different users and devices. A good classification can be achieved using a neural network with the *Reduced Feature Vector*. Providing a reasonably well basic functionality in the starting phase, with data collected on the device, the classification can be improved during usage.

For the further model application evaluation, the *Reduced Feature Vector NN 3* model was selected.

V. IMPROVED VIDEO STREAMING WITH SEAMLESS MPTCP HANDOVERS

As presented in Section IV, the learned neural network models reliably predict Wi-Fi connection loss with a high precision and recall. To show the usefulness of these results, we evaluated the performance when performing handovers in real-world mobile usage scenarios. In the following, we present a seamless Wi-Fi/cellular handover during DASH video streaming sessions.

A. Seamless Network Connectivity App

To gather the training data, perform the prediction, and test the applicability of the approach, we implemented a mobile application that performs the following tasks:

Sensor data collection and preprocessing: The sensor readings described in Section III-A are cached in memory and written periodically to a local SQLite database on the smartphone. When a run ends, the database is uploaded to a server. To execute the neural network on the smartphone, the sensor values are preprocessed similarly to the offline learning process. The mean, variance, and the observation window determined offline are used.

Online prediction: The offline learned neural networks are transpiled to Java using the sklearn-porter[12] framework, which allows execution of the same neural networks trained with sklearn on the device. This execution on the Android device allows us to achieve low delays in predictions, independence of Internet access, and protects user privacy.

Demonstration & reporting: We demonstrate the feasibility of the proposed approach using an embedded DASH video playback functionality. Here, the goal is to highlight potential in improved playback quality and stability made

possible by seamless connectivity. We use the open movie Elephants Dream, streamed from a server in the university network. The video was pre-encoded using the h.264 encoder for video and AAC for audio, in three bandwidths 1, 2 and 4 MBit/s and a segment length of 2 seconds. For video playback, the JavaScript-based DASH.js player (v 2.5.0) was used with a buffer size of 10 seconds in conjunction with the BOLA adaptation algorithm. To analyze the QoE, we collect and report raw video metrics in each streaming session while the video is playing to the server including stalls, playback bit rates, quality adaptations, and buffer levels for later evaluation.

MPTCP handovers: We use the Wi-Fi connection loss prediction to trigger the cellular subflow establishment for MPTCP *before* the Wi-Fi connection is lost. We implemented our approach on top of the MPTCP kernel implementation for Android² and the *MultipathControl*³ app of De Coninck et al. [3]. Furthermore, the video server uses MPTCP version 0.92 with the redundant scheduler and the fullmesh path manager enabled.

B. Experimental Setup

Our experiments consist of 3 connectivity modes in 4 scenarios each performed 5 times, resulting in a total of 60 iterations. The experiments were performed on a Google Nexus 5 smartphone running a rooted Android and the MPTCP Kernel version 0.89.5. The following connectivity modes were evaluated:

- *Stock Android:* The default Android mechanism was used to detect Wi-Fi unavailability. During these tests, no transition mechanism was used to have a baseline to compare with.
- *MPTCP:* To see how MPTCP can improve handover situations, it was enabled for the entire run in these tests. The cellular uplink was used as the second interface, and both client and server used the default scheduler.
- *Seamless:* During these tests, the *Reduced Feature Vector* neural network in configuration *NN 3* was used, since it showed the most promising results. MPTCP is enabled when a Wi-Fi loss is predicted and disabled when Wi-Fi is available and no loss is predicted for 5 seconds.

The following set of routes is chosen to evaluate scenarios in which Wi-Fi connection losses can occur. Figure 3 shows the room plan of the university building where the tests were performed.

Scenario 1: Leaving the office: Starting in the office, the smartphone is connected to the office Wi-Fi. The tester leaves after 120 seconds of video playback and heads towards the exit of the building. After the Wi-Fi connection is lost (determined in advance, roughly 50 meters) the tester waits for 10 seconds and ends the scenario.

Scenario 2: Visiting a colleague: The beginning is similar to *Scenario 1*, but the tester walks around about 20 meters away from the office, visiting a colleague, but not leaving the

²<https://multipath-tcp.org/pmwiki.php/Users/Android>

³<https://github.com/MPTCP-smartphone-thesis/MultipathControl>

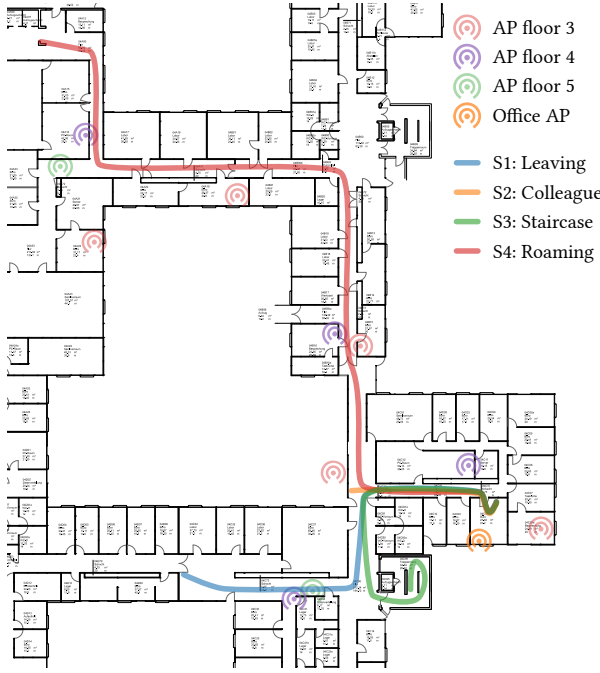


Figure 3: Map with Wi-Fi APs and scenarios routes.

Wi-Fi range. The tester stays for 10 seconds and then walks back to the office.

Scenario 3: Using the staircase: Starting as before, the tester leaves the office on the same route, but then uses the staircase to go up one floor and stays there for 10 seconds. The scenario shows the impact of a Wi-Fi connection that, while remaining available, is not usable.

Scenario 4: Wi-Fi roaming support: Starting in the office, the device is connected to the university network. The tester leaves after 120 seconds and heads towards the other end of the building, roaming between multiple possible Wi-Fi APs shown in Figure 3. The tester stays near the exit for 10 seconds and then walks back the same route. This scenario is created to further investigate the support of roaming gaps in corporate wireless networks where roaming might be available, but is not sufficient to achieve a high QoE.

C. Measuring Quality of Experience

The DASH video streaming technology is widely available, used by many vendors, and evaluated well. To measure the perceived QoE, several technical values are captured that are used to compute mean opinion scores (MOS) [20], as discussed below.

1) *Direct Measurements:* During the experiments, the DASH video player reports different technical parameters to a server. At the beginning of a video, the initial buffer has to be filled. This takes time, resulting in *initial stallings* that are perceived to be more disturbing the longer they take. Furthermore, *stalling events* during the video are also reported. Apart from the stallings, the *number of adaptations* is counted, since many adaptations also negatively influence the QoE. In addition, the *percentage of time spent in the highest*

achieved quality is measured. From a user's perspective, it is better to hold a certain quality as long as possible, even if it is not the best quality available. Since stalling events and quality adaptations partly depend on the buffer level (i.e., how much playable video is in the buffer), the *buffer level* is also captured. The buffer level should be as constant as possible for about 10 seconds.

Finally, a packet dump is performed on the server to allow further analysis of the connections created by MPTCP.

2) *QoE Metrics:* Apart from directly evaluating the metrics discussed above, derived metrics are used to capture relations between these metrics and their impact on QoE. The QoE_{stall} (Equation (1)) is derived on a MOS scale (where 1 denotes a bad user experience and 5 an excellent one) based on the stalling durations and frequencies during video playback. Furthermore, $MOS_{quality}$ (Equation (2)) is deduced based on playtime in the highest achieved quality (t). L denotes the average length of all conducted stallings (initially or during video playback) and N the number of stallings, again either initially or during playback. Since our work focuses on Wi-Fi connection loss events, we do not evaluate initial stallings.

$$MOS_{stall} = 3.5 \times e^{-(0.15 \times L + 0.19) \times N} + 1.5 \quad (1)$$

$$MOS_{quality} = 0.003 \times e^{0.064 \times t \times 100} + 2.498 \quad (2)$$

$$MOS_{combined} = \frac{MOS_{stall} + MOS_{quality}}{2} \quad (3)$$

Finally, Stohr et al. [20] propose the average MOS, denoted as $MOS_{combined}$ (Equation (3)), denoting a total user perception not only depending on stalling or quality adaptations. We use $MOS_{combined}$ to evaluate QoE.

3) *QoE Experimental Results:* In Table II, the overall results of the performed tests are presented, namely the number of stalling events (# St.) and the average duration of a stalling event (\emptyset St.), the number of adaptations (# A.), the relative time in the highest playback quality (HQ), and the average transmitted data (\emptyset TD).

Scenario 1: As shown in Table IIa, the *Stock* tests performed worst with 3 stalling events in total and an average stalling duration of about 1.5 seconds, while neither *MPTCP* nor *Seamless* tests did show any stalling events, which is a significant improvement compared to the stock tests. The amount of transferred data over cellular is high in the *MPTCP* test and low in the *Stock* test. *Seamless* results are between these two tests, thus saving cellular data compared to MPTCP, while still avoiding stallings. The results of these tests show that our prediction can avoid the handover gap completely.

When looking at the buffer levels, video stream quality and the used bandwidth, it can be seen that based on the prediction of *Seamless*, the cellular subflow is established proactively, resulting in a seamless handover and thus no video stalling.

Apart from improvements of these technical values, our approach improves QoE for users, as expressed in the $MOS_{combined}$. Figure 4 shows the $MOS_{combined}$ on the y-axis and the different connectivity modes on the x-axis, grouped by scenario. For the stock tests, the $MOS_{combined}$

Table II: Overview of Experimental Results

(a) Scenario 1: Leaving

Mode	# St.	\varnothing St.	# A.	HQ	\varnothing TD
<i>Stock</i>	3	1.46 s	23	87 %	21.75 MB
<i>MPTCP</i>	0	0 s	20	89 %	41.32 MB
<i>Seamless</i>	0	0 s	27	88 %	36.11 MB

(c) Scenario 3: Staircase

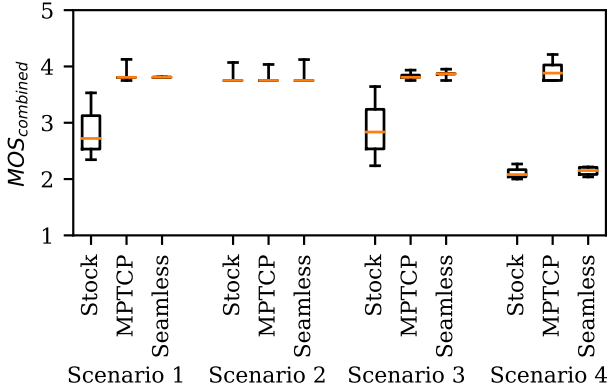
Mode	# St.	\varnothing St.	# A.	HQ	\varnothing TD
<i>Stock</i>	3	2.06 s	49	80 %	0 MB
<i>MPTCP</i>	0	0 s	32	87 %	33.92 MB
<i>Seamless</i>	0	0 s	28	85 %	16.81 MB

(b) Scenario 2: Colleague

Mode	# St.	\varnothing St.	# A.	HQ	\varnothing TD
<i>Stock</i>	0	0 s	10	92 %	0 MB
<i>MPTCP</i>	0	0 s	10	91 %	9.98 MB
<i>Seamless</i>	0	0 s	17	92 %	9.59 MB

(d) Scenario 4: Wi-Fi Roaming

Mode	# St.	\varnothing St.	# A.	HQ	\varnothing TD
<i>Stock</i>	18	14.98 s	42	53 %	0.89 MB
<i>MPTCP</i>	0	0 s	38	86 %	71.99 MB
<i>Seamless</i>	15	5.47 s	23	84 %	15.50 MB

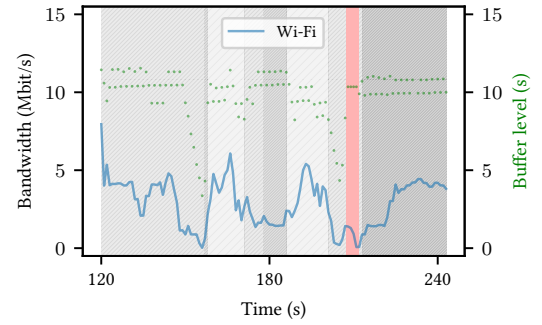
Figure 4: $MOS_{combined}$ values grouped to connectivity modes and scenarios.

is between about 2.5 (poor) and 3.5 (fair), indicating that the playback is not totally unsatisfactory, but far away from a great experience. *Seamless*, on the other hand, achieves a $MOS_{combined}$ of almost 4, indicating a good QoE, as high as in *MPTCP* tests.

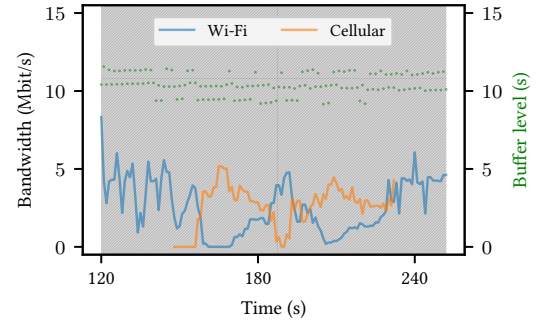
Scenario 2: As shown in Table IIb, all tests are comparable for all metrics, showing that our approach does not introduce any negative effects in already good situations. The transferred amount of data over cellular in *Seamless* is about as high as in the *MPTCP* tests. This is because the classifier predicts a Wi-Fi connection loss due to the movement of the smartphone and thus switches to the cellular network, even though this is not necessary.

Neither the technical metrics like buffer level or used bandwidth, nor the MOS values differ in these experiments, thus they are not further evaluated here, again indicating that our approach does not worsen the situation by any means.

Scenario 3: As shown in Table IIc, the stock tests performed worst with 3 stallings and an average stalling time of about 2 seconds. Additionally, with 49 adaptations and only 80% of the time at the highest achieved quality, the stock tests perform badly. *MPTCP* and *Seamless* do not stall at all. With 28 and 30 adaptations and 85% of the time at the highest achieved quality, the results of our approach are as good as in the *MPTCP* tests, again showing significant improvements



(a) Stock Android



(b) Seamless

Figure 5: *Stock* and *Seamless* in *Scenario 3*

over the stock implementation. The data usage over cellular shows the same behavior as in *Scenario 1*.

Figures 5a and 5b show bandwidth and buffer level for *Scenario 3*. In the stock tests, the maximum distance is shown in the used bandwidth around seconds 150 and 210. *Seamless* improves this situation and establishes a cellular connection in a timely manner resulting in no stallings. $MOS_{combined}$ during the stock tests shows again a relatively bad QoE with about 2.5 to 3.5 compared to the high MOS values of about 4 during *MPTCP* and tests using our approach.

Scenario 4: The results of the stock tests in Table II d indicate that Android does not handle *Scenario 4* well. The video stalls 18 times and for about 15 seconds on average. The video quality adapts 42 times in total and stays only for 53% of the time at the highest achieved quality. *MPTCP*, on the other hand, handles *Scenario 4* very well with no stallings,

few quality adaptations, and 86% of the time at the highest achieved quality.

Although *Seamless* cannot completely cope with the situation, the results are much better than in the stock tests. With 15 stallings and an average stalling duration of 5.5 seconds, just 23 quality adaptations and 84% percent of the playtime at the highest achievable quality, the results indicate an improved QoE using our approach. The benefits of these improvements come with the cost of using more data (14.61 MB) over the cellular network, but only 21.53% of cellular data compared to the MPTCP tests.

It is evident that our approach predicts Wi-Fi connection loss correctly, since connection establishment occurs in a timely manner. However, the cellular interface does not reach the high bandwidth used by MPTCP in the same scenario, which might be due to the fact that the cellular interface requires a longer starting phase in the concrete area of the building. Also, due to the relatively short Wi-Fi-less gaps investigated in this scenario, the cellular connections are dismantled shortly after they are established. A model optimized not only for predicting Wi-Fi connection loss events but also Wi-Fi recovery could improve such scenarios by keeping the cellular link longer alive.

$MOS_{combined}$ during the *Seamless* and stock tests is comparably bad with a value of about 2, while MPTCP still reaches a MOS of about 4. Nevertheless, our approach reaches a slightly higher QoE than stock Android, as shown in Figure 4.

VI. CONCLUSION

In this paper, we proposed a novel data-driven approach to predict Wi-Fi connection loss to perform seamless vertical Wi-Fi/cellular handovers. The approach is based on sensors available in today's smartphones and uses MPTCP to dynamically switch between different wireless connectivity modes. We demonstrated that our trained neural networks reliably predict Wi-Fi connection loss 15 seconds ahead of time when users move around, with a precision of up to 0.97 and a recall of up to 0.98. Furthermore, we illustrated the benefits of our Wi-Fi connection loss prediction approach with an MPTCP video streaming application. We showed that our predictions improve the QoE mean opinion score from 2.7 to up to 3.8 for certain scenarios, while reducing the required cellular data usage by up to 50% compared to traditional MPTCP approaches, with a negligible power consumption overhead.

There are several areas for future work. While the contextual sensors used in our work support high-quality predictions, other more domain-specific sensors might be useful to predict, e.g., Wi-Fi overloads. It would also be interesting to learn predictions for user/access point combinations. To deploy predictors efficiently on off-the-shelf smartphones, lightweight neural networks on dedicated processing engines should be considered. Finally, Wi-Fi connection regain prediction is an interesting area for future research.

ACKNOWLEDGMENT

This work has been funded by the German Research Foundation (DFG, SFB 1053 MAKI).

REFERENCES

- [1] A. Ahmed, L. M. Boulahia, and D. Gaiti, "Enabling vertical handover decisions in heterogeneous wireless networks: A state-of-the-art and a classification," *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 776–811, 2014.
- [2] Y.-C. Chen, Y.-s. Lim, R. J. Gibbens, E. M. Nahum, R. Khalili, and D. Towsley, "A measurement-based study of multipath TCP performance over wireless networks," in *Internet Measurement Conference*. ACM, 2013.
- [3] Q. De Coninck, M. Baerts, B. Hesmans, and O. Bonaventure, "A first analysis of multipath TCP on smartphones," in *17th Int. Passive and Active Measurements Conference*, vol. 17. Springer, 2016.
- [4] Q. De Coninck and O. Bonaventure, "Every millisecond counts: Tuning multipath TCP for interactive applications on smartphones," Technical report. Available at <http://hdl.handle.net/2078.1/185717>, Tech. Rep., 2017.
- [5] A. Ford, C. Raiciu, M. Handley, and O. Bonaventure, "TCP extensions for multipath operation with multiple addresses," RFC 6824, Internet Engineering Task Force, 2013.
- [6] S. Horrich, S. B. Jamaa, and P. Godlewski, "Adaptive vertical mobility decision in heterogeneous networks," in *3rd Int. Conf. on Wireless and Mobile Communications*, March 2007, pp. 44–44.
- [7] M. Khan, A. Ahmad, S. Khalid, S. H. Ahmed, S. Jabbar, and J. Ahmad, "Fuzzy based multi-criteria vertical handover decision modeling in heterogeneous wireless networks," *Multimedia Tools and Applications*, vol. 76, no. 23, pp. 24 649–24 674, 2017.
- [8] A. Liaw and M. Wiener, "Classification and regression by RandomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
- [9] T. Lin, C. Wang, and P.-C. Lin, "A neural-network-based context-aware handoff algorithm for multimedia computing," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 4, no. 3, pp. 17:1–17:23, Sep. 2008.
- [10] L. Ma, F. Yu, V. C. Leung, and T. Randhawa, "A new method to support UMTS/WLAN vertical handover using SCTP," *IEEE Wireless Communications*, vol. 11, no. 4, pp. 44–51, 2004.
- [11] A. A. Mansour, N. Enneya, and M. Ouadou, "A seamless handover based MIH-assisted PMIPv6 in heterogeneous network (LTE-WIFI)," in *2nd Int. Conf. on Big Data, Cloud and Applications*. ACM, 2017, pp. 67:1–67:5.
- [12] D. Morawiec, "sklearn-porter," Transpile trained Scikit-learn estimators to C, Java, JavaScript and others. [Online]. Available: <https://github.com/nok/sklearn-porter>
- [13] N. Nasser, S. Guizani, and E. Al-Masri, "Middleware vertical handoff manager: A neural network-based solution," in *2007 IEEE International Conference on Communications*, June 2007, pp. 5671–5676.
- [14] A. J. Nicholson and B. D. Noble, "Breadcrumbs: Forecasting mobile connectivity," in *14th ACM Int. Conf. on Mobile Computing and Networking*, ser. MobiCom '08. ACM, 2008, pp. 46–57.
- [15] C. Paasch, G. Detal, F. Duchene, C. Raiciu, and O. Bonaventure, "Exploring mobile/wifi handover with multipath TCP," in *ACM SIGCOMM Workshop Cellnet*, 2012.
- [16] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, no. Oct, pp. 2825–2830, 2011. [Online]. Available: <http://scikit-learn.org/>
- [17] C. E. Perkins, "IP mobility support for IPv4, revised," RFC 5944, Internet Engineering Task Force, 2010.
- [18] C. Pluntke, L. Eggert, and N. Kiukkonen, "Saving mobile device energy with multipath TCP," in *6th International Workshop on MobiArch*. ACM, 2011, pp. 1–6.
- [19] C. Raiciu, C. Paasch, S. Barre, A. Ford, M. Honda, F. Duchene, O. Bonaventure, M. Handley *et al.*, "How hard can it be? designing and implementing a deployable multipath TCP," in *NSDI*, 2012. [Online]. Available: <https://dl.acm.org/citation.cfm?id=2228338>
- [20] D. Stohr, A. Frömmgen, J. Fornoff, M. Zink, A. Buchmann, and W. Effelsberg, "QoE analysis of DASH cross-layer dependencies by extensive network emulation," in *2016 Workshop on QoE-based Analysis and Management of Data Communication Networks*. ACM, 2016, pp. 25–30.
- [21] W. Wanalertlak, B. Lee, C. Yu, M. Kim, S.-M. Park, and W.-T. Kim, "Behavior-based mobility prediction for seamless handoffs in mobile wireless networks," *Wireless Networks*, vol. 17, no. 3, pp. 645–658, 2011.