Oğuz and Kalaycıoğlu
J Wireless Com Network (2022) 2022:101
https://doi.org/10.1186/s13638-022-02183-7

EURASIP Journal on Wireless Communications and Networking

RESEARCH Open Access

Anomaly detection in multi-tiered cellular networks using LSTM and 1D CNN

Hasan Tahsin Oğuz* and Aykut Kalaycıoğlu

*Correspondence:
htoguz@ankara.edu.tr

Department of Electrical and Electronics Engineering, Ankara University, Ankara, Turkey

Abstract

Self-organizing networks (SONs) are considered as one of the key features for automation of network management in new generation of mobile communications. The upcoming fifth-generation mobile networks and beyond are likely to offer new advancements for SON solutions. In SON concept, self-healing is a prominent task which comes along with cell outage detection and cell outage compensation. Next-generation cellular networks are supposed to have ultra-dense deployments which make cell outage detection critical and harder for network maintenance. Therefore, by imitating the ultra-dense multi-tiered scenarios, this study scrutinizes femtocell outage detection with the help of long short-term memory and one-dimensional convolutional neural networks by using time sequences of key performance indicator parameters generated in user equipment. In both the proposed schemes, probable outage-related anomalies in femto access points (FAP) are detected and classified within predetermined time sequence intervals. Moreover, aggregation decision methods are also incorporated into the proposed framework for boosting cell outage detection procedure on FAP level. Our findings show that proposed deep learning approaches outperform existing feed-forward neural networks, and on the average, in more than 80% of the cases the outage states of the femtocells are correctly predicted among healthy and three anomalous states.

Keywords: Cell outage detection, Self-organizing networks, LSTM, Femtocells, 1D CNN, Deep learning, FFNN

1 Introduction

Next-generation cellular systems are likely to set forth a lot of advancements in mobile communications such as more capacity, higher rates, ultra-low latency, massive connectivity, and lower energy consumption. Ultra-dense architectures are supposed to be one of the major features of these improvements. In this regard, for better network maintenance, Third-Generation Partnership Project (3GPP) has already come up with the concept of self-organizing networks (SON) which involves self-configuration, self-optimization, and self-healing [1]. In modern large-scale cellular network architectures, management of SONs containing tremendous number of network nodes becomes very complicated under the existence of anomalies which have the potential to degrade the user's quality of experience (QoE). Anomalies may arise from various factors such as

unusual traffic conditions, illegitimate intrusions, misconfiguration, security threats as well as outages and cell degradation [2–4]. Therefore, efficient detection of anomalies is critical for sustainability and maintenance in modern cellular networks.

This study focuses on anomalies arising from cell outage and cell degradation among the types of mentioned anomalous conditions in cellular networks. Cell outage is defined as a state of the base station where all or most of the user equipment (UE) in the cell are unable to establish or keep its radio connectivity with the result of degradation in capacity and coverage gaps [1, 5]. Cell outage detection (COD) is considered as the first primary step for self-healing capability of the modern SONs. Hardware, software, or external failures such as power cut or network disconnection are the main causes for outages.

COD is basically a classification problem about detecting the non-healthy cell among healthy ones by making use of some statistics generated by UEs, base stations, and other network components. Traditionally, besides user complaints, detection of outages may involve site visits, drive tests, and manual analysis of the alarms generated on the operations support systems (OSS), which make outage detection costly. Future mobile networks are supposed to handle COD autonomously, such that detection algorithms employed on the OSS should continuously keep track of UE statistics and process the data generated on the UEs as suggested by 3GPP releases. Especially in the existence of multi-tiered dense deployments, outages regarding small cells may not be detected for a long time. Unlike macro-cells, small cells such as femtocells, picocells, and microcells introduce extra challenges and complexities to the system. Therefore, outage detection on small cells is regarded as a much harder task due to sparse user statistics and vertical handovers [6–9], thus justifying the consideration of more advanced methods.

Cell outage detection studies mainly focus on macro-cell anomalies. Within this scope, handover statistics based on KPI measures are employed on COD analysis in [10]. In [11], COD is handled with the help of neighbor cell list reports by detecting outage cells according to the changes in the topology generated by visibility graphs. In [12], channel quality indicator (CQI) is used within a composite hypothesis for outage detection by means of a discriminant function. Machine learning methods are also popular in outage detection of macro-cells. In [13] and [14], clustering algorithms and Bayesian networks are conducted for COD, respectively. In [15], alternative to machine learning procedures, an anomaly detection method based on statistical processing of big data emerged from KPI measures is introduced.

Some researchers propose approaches on detecting cell outages regardless of the type of the anomalous base stations in multi-tiered architectures containing macro, micro, femto or pico base stations. Within this context, K-nearest neighbors method is conducted for COD in multi-tiered networks [7]. Hidden Markov model (HMM), another well-known maximum likelihood classifier, is also studied on COD by training KPI parameters regarding healthy cells and outage cells for predicting the outage status of the base stations [5]. A method, based on the analysis of time evolution metrics, is proposed in [16] for detecting faulty patterns arising from degraded cells. In another study, researchers make use of cell traces along with big data analytic techniques and apply aggregation at the cell levels for identifying any problematic cases in cellular networks [17]. Besides, a classification tree-based method along with cell-level aggregation for

identification of faulty cells is proposed in [18]. On the other hand, some researchers present studies about COD in only small cells so as to focus on its difficulties when compared to relatively easier detection of macro-cell anomalies. In this context, Wang et al. propose a cooperative COD scheme for femto cells in a conventional heterogeneous network by means of spatial correlations among users [6].

The 5G literature regarding anomaly detection studies recently gets attention from many researchers. Intrusions, abnormal traffic loads, security threats, and malicious network traffic on the 5G networks are discussed in the latest outstanding studies. In this respect, in [19], researchers propose a deep learning-based 5G-oriented cyberdefence architecture. In [2], an unsupervised learning scheme is proposed for anomaly detection in mobile networks using real CDR data. In another study, authors employ a learning-based anomaly detection framework using service logs and query traces [20]. In [21], researchers present detection of phishing attacks in the network by making use of a triple-stage privacy-aware classification framework. In [3], the authors study identification of traffic types with different machine learning tools by analyzing network traffic data and they identify malicious network patterns. In another study, researchers propose an auto-encoder-based anomaly detection method for Internet of things (IoT) networks such that the proposed architecture is deployed in both IoT devices and mobile core network facility [22]. The authors also propose a deep semi-supervised anomaly detection scheme for identifying legitimate activities that may yield abnormal traffic like concerts and football games [4]. In this study, they use mobile networks as a supplementary sensing platform for detecting urban anomalies.

Anomaly detection in terms of cell outage and cell degradation, which is the main goal of this study, is also recently covered in the 5G literature. Within this context, in [23] the researchers study outage detection on ultra-dense heterogeneous networks for different shadowing conditions by making use of entropy field decomposition method. Another study focuses on an unsupervised learning for anomaly detection in mobile networks using call detail records (CDR) data and distinct traffic patterns [24]. Recently, deep learning approaches have started to gain interest among researchers in the area of COD. In [25], recurrent neural networks are comparatively analyzed along with traditional support vector machines in terms of COD performance. Another study is introduced in [26] where anomaly detection in cellular networks is studied by using convolutional neural networks (CNN).

This study aims to develop a foresight for cell outage management in existing cellular systems and forthcoming next-generation networks within multi-tiered ultra-dense deployments. Inspired by the advancements of deep learning methods on time sequence analysis, we propose using long short-term memory and one-dimensional CNN, concomitantly with aggregation decision methods, for detection of femtocell outages which pose extra challenges. In the proposed scheme, we suggest not only detecting femtocell outages but also adroitly classifying the type of anomalies into three outage subclasses according to the severity of the degradation of the femto access point functionality.

As a variant of recurrent neural networks (RNNs), LSTM modules have been used with time sequence labeling tasks on many areas so far. Similarly, 1D CNN structures have also been used for extracting features from fixed-length data like audio recordings and various other time series of sensor data. Thereby, in this framework,

we separately employ LSTM and 1D CNN for the investigation of outage patterns by using time sequences of metrics measured on UEs placed around the femtocell sites. Signal-to-interference plus noise ratio (SINR) and CQI are utilized as input feature time series data within the aforementioned deep learning structures for training and testing phases.

The outstanding contributions of this study can be summarized as follows:

(1) This study demonstrates two deep network approaches, namely LSTM and 1D CNN, for detection and identification of anomalous states of densely deployed femtocells to inspire outage management in existing and upcoming cellular networks.
(2) This study introduces aggregation decision methods integrated with LSTM and 1D CNN for boosting the cell outage detection performance.
(3) This study shows that proposed 1D CNN and LSTM approaches outperform existing feed-forward neural networks (FFNN) on cell outage detection in ultra-dense cellular networks.

The rest of the paper is organized as follows. Section 2 presents employed radio access network (RAN) structure and the deep learning methods. Section 3 introduces the details of the outage detection, classification algorithms, training procedures, and aggregation. Section 4 shows results and includes essential hermeneutic discussion issues. Finally, Sect. 5 concludes the paper.

2 Methods/experimental

This study investigates the detection of anomalous states of FAPs in ultra-dense deployments for providing an insight into the management of self-organizing next-generation cellular networks. In this study, femtocells of a cellular system are considered to be in one of the four states as healthy, degraded, crippled, and catatonic states such that the service quality of FAPs falls into a reduction in non-healthy states [6].

Degraded FAPs have a slightly lower performance than healthy ones and can resume normal operation after the environmental effects causing the anomaly disappear. Crippled FAPs have serious problems and may carry very little traffic. On the other hand, catatonic FAPs are generally out of service due to catastrophic failures like serious power cuts [5]. When there is a reduction in FAP's output power due to hardware or software failures such as implementation failures in channel processing, external power supply problems or even misconfiguration, FAPs undergo anomalous states from the healthy state. However, service providers generally may not realize these types of state changes very quickly and efficiently. Making use of the measured UE data related to received signal strength, which are also reported to base stations, the anomalous states might be detected by the use of deep network applications like LSTM and 1D CNN which have been employed in many fields so far. LSTM is utilized in this study for its ability to learn time interactions on UEs' time series data to detect anomalies regarding outages. On the other hand, another deep learning method, 1D CNN structure, is also employed for its powerful extraction capability of spatiotemporal anomaly patterns that might be generated in outage events.

2.1 Key performance indicators

KPIs are generally utilized for monitoring and optimizing cellular network performance. Therefore, KPIs may also be well suited to anomaly detection tasks. KPI metrics such as CQI, reference signal received power (RSRP), reference signal received quality (RSRQ), CQI, and SINR are all candidates that may supply clues and information intended for detection of anomalies and monitoring the healthiness of the network. Among these, our detection algorithms are confined to two metrics, CQI and SINR, for their relatively lower computation cost. CQI carries information about the channel quality of the communication link, and SINR is basically the signal-to-interference plus noise ratio measured on a UE.

KPI metrics of every UE are reported to OSS via base stations. In the OSS, CQI and SINR values of every UE are to be monitored for a certain time interval of T for our proposed detection schemes. Two featured KPI data structure of UEs time sequence employed in our study, for any user (e.g., UE_n), is shown in Eq. (1):

$$\mathrm{user}_n = \begin{pmatrix} cqi_1 & cqi_2 & \dots & cqi_{t-1} & cqi_t & cqi_{t+1} & \dots & cqi_{\mathrm{T}} \\ \mathrm{sin}r_1 & \mathrm{sin}r_2 & \dots & \mathrm{sin}r_{t-1} & \mathrm{sin}r_t & \mathrm{sin}r_{t+1} & \dots & \mathrm{sin}r_{\mathrm{T}} \end{pmatrix} \tag{1}$$

2.2 Radio access network

In this study, we build radio access network structure with the help of a well-known downlink system level simulator for cellular networks [27]. In order to synthesize time sequences of UE data, we employ macro BSs with tri-sectored structure in the hexagonal geometry using 7 macro BSs with each macro BS having 120 UEs. Additionally, we generate 2 FAPs in each sector of every macro-cell cluster and 5 UEs in each femtocell. The positions of the UEs are randomly generated. The FAPs operate in open subscriber group configuration in which the femtocells also service users from other base stations. The parameters of the system level simulations are summarized in Table 1. We adjust the output power of the FAPs as described in [5] to generate the abnormal conditions. In our proposed framework, we employ healthy, degraded, crippled, and catatonic FAPs that radiate at 30 dBm, 20 dBm, 10 dBm, and −10 dBm, respectively. We run the simulations for a duration of 60 milliseconds and record UE data in every millisecond. Within this duration, for imitating the anomalous cases, we degrade the transmit power of the FAPs in a random time for all three possible anomalous cases with the condition that FAPs are initially in healthy state. In our work, we employ three different shadowing conditions with standard deviations 2, 5, and 8 dB to scrutinize the fading effects on outage detection.

2.3 Long short-term memory

In deep learning field, LSTM is known to be a special type of artificial RNN architecture. RNNs have the ability of incorporating activations from previous time steps as inputs to make decisions for the current output and current state. This makes RNNs efficient for time sequence classification purposes; nevertheless, RNNs suffer the bottleneck of gradient vanishing problem. At this point, LSTM comes out with a memory cell unit and four gates layers which interact and cooperate in a special structure so as

Table 1 RAN simulation parameters

Parameters	Values
Cellular layout	7 macro-cells (three-sectored) in hexagonal geometry A total of 42 femtocells (2 FAPs per sector)
Macro BS Tx power	46 dBm
Femto BS Tx power for healthy case	30 dBm
Femto BS Tx power for degraded case	20 dBm
Femto BS Tx power for crippled case	10 dBm
Femto BS Tx power for catatonic case	−10 dBm
Path loss model	TS36942 urban
Number of UEs per macro-cell	120
Number of UEs per sector	40
Number of UEs per femtocell	5 UEs
Inter eNodeB distance	500 m
Channel bandwidth	20 MHz
User location and mobility	Uniformly distributed locations with random walk model

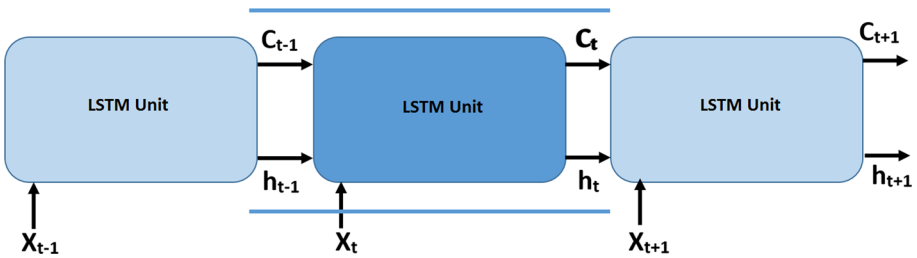

Fig. 1 Repetitive architecture of LSTM units in time steps

to eliminate gradient vanishing. LSTM networks are useful in classification and processing of time series. Besides, LSTM can be suitable for making predictions on time sequences since it has the ability to learn long-term relationships. It can recall information for past periods of time, and it can cope with lags between critical points in time series. Within this framework, in [28] researchers made use of LSTM for anomaly detection in cellular networks.

Unlike RNNs having one hyperbolic tangent layer, a typical LSTM unit is composed of a cell being the memory part, an input gate, an input modulation gate, an output gate, and a forget gate. Input gate decides the entrance of a new data to the cell, whereas input modulation gate controls the extent to which the new data enter the cell. Forget gate decides what information will be eliminated from the cell state, and the output gate controls the extent to which the value in the cell is used to contribute to the activation of the LSTM unit output.

Recurrent neural networks, including LSTM networks, consist of repetitive sequences of neural network modules in chain arrangement as shown in Fig. 1. By use of activations from previous cycles as inputs to the current network, a decision for the current input can be made, so that LSTM can better be suited for sequential labeling purposes [29, 30]. Correspondingly, LSTM units can be trained with training sequence sets in a supervised manner by means of an optimization algorithm. For the optimization process, gradient descent method with back-propagation in time can be used to compute the required gradients. By this way, each weight of the LSTM network changes in proportion to the derivative of the error at the output layer of the LSTM unit with respect to corresponding weight. The relations concerning the hidden state h_{t-1}, current state c_t, and input of the modules x_t are expressed in a set of equations; $f_t{=}\sigma\,(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$, $i_t = \sigma\ (W_{xi}x_t + W_{hi}h_{t-1} + b_i)$, $\check{c}_t{=}\tanh\,(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$, $c_t{=}f_t \circ c_{t-1} + i_t \circ \check{c}_t$, $o_t = \sigma\,(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$, and $h_t{=}o_t \circ \tanh c_t$. In these compact equations, f_t is the forget gate value, i_t is the input gate value, $\check{c}_t$ is the input modulation gate value, o_t is the output gate value, c_t is the current state (cell memory) value, h_t is the value of the output (hidden state) of the LSTM unit at time step t, σ is the activation function (sigmoid, ReLU, etc.), tanh is the hyperbolic tangent function, and $\circ$ is the element-wise Hadamard product. W_{xf}, W_{hf}, W_{hi}, W_{xi}, W_{xc}, W_{hc} W_{xo}, W_{ho} are the weights of the inputs and the recurrent connections. W_{xf} is the weight from input to forget gate, W_{hf} is the weight from previous hidden state to forget gate, W_{hi} is the weight from previous hidden state to input gate, W_{xi} is the weight from input to input gate, W_{xc} is the weight from input to input modulation gate, W_{hc} is the weight from previous hidden state to input modulation gate, W_{xo} is the weight from output to output gate, and W_{ho} is the weight from previous hidden state to output gate. Additionally, b_f, b_i, b_c, and b_o are the bias values for the forget, input, input modulation, and output gates, respectively. The detailed internal structure of a single LSTM unit is shown in Fig. 2.

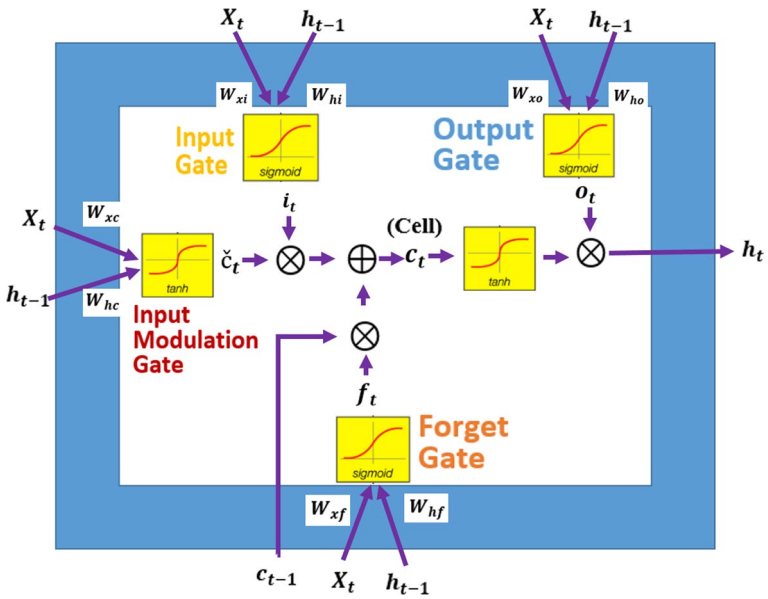

Fig. 2 Unit structure of LSTM

Finally, in our envisaged LSTM model, besides our input of size 2x60, we employ a mini-batch size of 27, hidden neuron units of 50 and a maximum epoch number of 100 for training. Sequentially, we construct an input layer, an LSTM layer like the one shown in Fig. 2 with 50 hidden neurons, 4 fully connected layers, a soft-max layer, and a final classification layer at the end.

2.4 One-dimensional convolutional neural networks

CNNs are very powerful tools in many modern artificial intelligence applications, particularly in machine learning and computer vision tasks. The architecture of a generic CNN has one or more convolutional layers, followed by a pooling layer, a flattening layer, and a fully connected layer. These layers help in learning the features, patterns, and objects in the data of interest.

In the literature, many CNN applications have multidimensional structure, especially in tasks related to image and video data [31] as well as intrusion detection tasks in networks [32, 33]. On the other hand, 1D CNN is also used in data having time sequence character such as text, handwriting, speech signals, and natural language processing [34]. 1D CNN searches for temporal patterns and differences in the direction of elapsing time via a convolution kernel window [35]. In our case, 1D CNN kernel of size 2 is slid toward the direction of elapsing time so that it can extract the temporal pattern changes in the time data composed of CQI and SINR values of subject UEs for investigating any probable anomalies.

The first layer, $conv_1$, in our 1D CNN structure is the first convolutional layer. This layer is fed with the input data volume, $I \in \mathbb{R}^{(n_{IH})\times(n_{IW})}$ and learning filters, also called kernels, $F \in \mathbb{R}^{(n_{FH})\times(n_{FW})}$ where n_{IH}, n_{FH} are height and width of input volume and n_{IW} and n_{FW} are the height and width of the filters applied, respectively. In this layer, parallel convolution processes are employed between each kernel and the input volume in the desired directions followed by a bias and a rectified linear unit function (ReLu) operation to generate the output of this layer, $Y_{\mathrm{conv1}} \in \mathbb{R}^{(n_{IH}-n_{FH}+1)\times(n_{IW})}$, which is also the input of the second convolutional layer ($conv_2$), I_{conv2}. The output of the second convolutional layer, $Y_{\mathrm{conv2}} \in \mathbb{R}^{(n_{IH}-2n_{FH}+2)\times(n_{IW})}$ is also the input to pooling layer, I_{maxpool}. As illustrated in Fig. 3, two subsequent convolutional layers are conducted in our CNN architecture for better extraction of features consisting of lower level features. At the end of convolutional layers, activation maps, also called as feature maps,

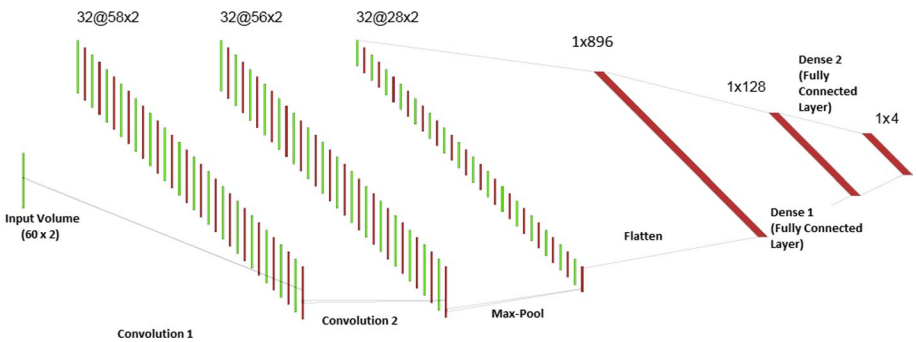

Fig. 3 Employed CNN layers with dimensions

holding the kernel responses for every spatial position are generated. In our work, the number of kernels N_{kernel} is 32 which means that 32 probable patterns are subject to our feature extraction task.

In our study, we set filter height, n_{FH} to 3 since the sign of anomalies in CQI and SINR of UEs generally appears suddenly within a few milliseconds. These sudden pattern deviations due to anomalous degradations are also the main reason why we do not apply any strides or zero paddings to our data to avoid missing any significant patterns. In our study, the height of our input volume n_{IH} is set to 60 which is the employed UE time sequence duration in milliseconds. Since we use time sequences of two metrics, width of our input volume, n_{IW} equals 2. Accordingly, the width of our filter, n_{FW} is also set to 2.

In the sequential structure of CNNs, pooling layers generally take place right after the convolutional layers to reduce the spatial size and parameters within a downsampling character. We employ a max-pooling filtering operation which computes the largest value in each patch of every activation map. The input to this layer I_{maxpool} is reduced according to the max-pooling filter size, n_{mp} which can be considered as the downsampling rate and is chosen as 2 in our study.

The output of the max-pooling layer, Y_{maxpool} with the given dimensions $\left\lfloor \frac{(n_{\text{IH}} - 2n_{\text{FH}} + 2)}{n_{\text{mp}}} \right\rfloor \times (n_{\text{IW}})$ is also the input of the flattening layer, I_{flat}. Flattening layer reshapes the input by converting it into a one-dimensional long feature vector output, Y_{flat} with dimensions $\left(\left\lfloor \frac{(n_{\text{IH}} - 2n_{\text{FH}} + 2)}{n_{\text{mp}}} \right\rfloor \times (N_{\text{kernels}}) \right) \times 1$, which is also the input of the last fully connected layer, I_{Dense}. Fully connected layers, following the flattening layer, basically emit the probability distribution of the classes that the overall net has to choose. In other words, the feature maps are transformed into output predictions for the model where each hidden unit is connected to all hidden units of the previous flattening layer. At the very end of the layered structure, we employ two consecutive fully connected layers where the first one has 128 and the second one has 4 neurons. The output of these two cascaded fully connected layers gives us the probability distribution of the classes corresponding to the input volume. The predicted class related to the input volume is decided according to most probable class.

3 Implementation of anomaly detection

Our proposed anomaly detection approaches involve continuous monitoring of the UE signals around the FAP of interest by the operations support system. Once our trained deep learning model senses an anomaly in the femtocell, an alarm is generated. Unlike macro BSs, FAPs are generally serving much fewer number of users which makes them harder to detect in the case of an anomalous situation due to lack of information emanating from sparse user characteristics on the site (i.e., 1–5 users per femtocell). For our anomaly detection analysis, we made use of the time sequence data of the UEs residing inside a circle of a certain radius centered at the FAP of interest. We call this radius for FAP region of interest as ROI in the rest of the paper. In this study, for UEs to be involved in the anomaly detection process we employ widely accepted typical values for ROI as 30, 40, and 50 meters [36, 37].

3.1 Model training

The first step of proposed anomaly detection scheme involves training of our deep network structure with two featured (SINR and CQI) time sequences of every UE in ROI for healthy and all three possible anomalous cases. Once training is complete, time sequences of the test UEs in ROI are tested on the trained model with tenfold cross-validation. In other words, two featured time sequence data of UEs are fed into trained model structure and a classification decision is made for the FAP. It is worth mentioning that in practical applications, training step is done for once at the beginning of network monitoring and hence will neither take time nor bring computational burden in classification and detection phases in real time.

3.2 Classification

Every test sequence of the UEs residing in ROI is applied to the trained model so as to produce a discrete probability distribution vector PD_{UE}, with four probability values as expressed in Eq. (2). For every UE in ROI, PD_{UE} vector holds p_{healthy}, p_{degraded}, p_{crippled}, and $p_{\mathrm{catatonic}}$ values. These values are the probabilities of a single UE being in a healthy, degraded, crippled, and catatonic cell, respectively. Thus, the probabilities in this vector are accepted as scores revealing the degree of involvement of each UE with the given anomalous or healthy states. The classifier predicts the anomalous state of the UE time sequence by choosing the class with the highest probability in PD_{UE}.

$$PD_{\mathrm{UE}} = [p_{\mathrm{healthy}}, p_{\mathrm{degraded}}, p_{\mathrm{crippled}}, p_{\mathrm{catatonic}}] \tag{2}$$

3.3 Aggregation decision

In this framework, we are motivated to make a decision about the state of the FAP rather than the state of the UE being exposed to such a probable anomaly. Therefore, we employ cell-level aggregation to predict the state of the FAP by making use of the scores of the UEs that form a cluster ensemble in ROI [17, 18, 38, 39]. Henceforth, once the probability scores of each state in PD_{UE} vector are computed for every UE in ROI, we employ an aggregation decision procedure based on PD_{UE} and come to a final decision regarding the state of the FAP of interest.

Assuming that the positions of the UEs and FAPs are known to the OSS, we simply compute the mean of the probabilities in PD_{UE} within all UEs residing in ROI. The class label with the highest average probability is simply decided as the state of the FAP. We call this aggregation procedure as *ensemble averaging* method.

Alternative to ensemble averaging, one can also propose another aggregation decision method, which we call *majority voting*. We start this procedure by examining the PD_{UE} of every UE in ROI, revealed by the classifier, and determine the predicted class of each UE according to maximum of the probabilities in PD_{UE} for every UE. Then, we check the majority among predicted class labels among UEs and continue as follows:

- In the case of equity in majority check among predicted classes of UEs, namely equal number of UEs are assigned to two or more different predicted classes, we compute

the mean of the probabilities in PD_{UE} of each UE residing in ROI. The class label with the highest average probability is decided as the state of the FAP as in ensemble averaging.

- If there is no equity in majority of the predicted classes, meaning that there is only one major predicted class, then the state of the FAP is directly decided as the class of majority of the predicted states of UEs in ROI.

Pseudo-algorithms for ensemble averaging and majority voting methods are given as Algorithms 1 and 2, respectively. The results regarding majority voting and ensemble averaging are mentioned in the Results and Discussion section.

Algorithm 1: Ensemble Averaging

Input: PD_{UE}: matrix consisting of discrete probability distribution vectors for all N number of UEs residing in ROI, size: Nx4 where, any row of the PD_{UE}, is a vector in the form $[p_{\mathrm{healthy}}, p_{\mathrm{degraded}}, p_{\mathrm{crippled}}, p_{\mathrm{catatonic}}]$ of size: 1x4

Output: Predicted State for the FAP of interest (**predClass** encoded as: $1 \rightarrow healthy$, $2 \rightarrow degraded$, $3 \rightarrow crippled$, $4 \rightarrow catatonic$

begin
  $predClass \leftarrow$ init empty predicted state variable;
  $MeanOfColumns \leftarrow$ init empty mean vector (1x4) for columns of PD_{UE} ;
  for $i = 1 : 4$ **do**
    $MeanOfColumns(i) = (\frac{1}{N})\sum_{j=1}^{N} PD_{UE}(j, i)$;
  end
  $predClass = \underset{i}{\operatorname{argmax}}\, MeanOfColumns(i)$;
  return $predClass$;
end

Algorithm 2: Majority Voting

Input: PD_{UE}: matrix consisting of discrete probability distribution vectors for all N number of UEs residing in ROI, size: Nx4 where, any row of the PD_{UE}, is a vector in the form [p_{healthy}, p_{degraded}, p_{crippled}, $p_{\text{catatonic}}$] of size: 1x4

Output: Predicted State for the FAP of interest (**predClass** encoded as: $1 \rightarrow healthy$, $2 \rightarrow degraded$, $3 \rightarrow crippled$, $4 \rightarrow catatonic$

begin
  $predClass \leftarrow$ init empty predicted state variable;
  $stateCnt \leftarrow$ init empty counter vector (1x4) for predicted states;
  $MeanOfColumns \leftarrow$ init empty mean vector (1x4) for columns of PD_{UE} ;
  for $i = 1 : N$ **do**
    $UE\_state = \underset{j}{\text{argmax}}\ PD_{UE}(i, j)$;
    $stateCnt(UE\_state) = stateCnt(UE\_state) + 1$;
  end
  if $\underset{i}{\text{argmax}}\ stateCnt(i)$ *is unique* **then**
    $predClass = \underset{i}{\text{argmax}}\ stateCnt(i)$;
  else
    for $i = 1 : 4$ **do**
      $MeanOfColumns(i) = (\frac{1}{N})\sum_{j=1}^{N} PD_{UE}(j, i)$;
    end
    $predClass = \underset{i}{\text{argmax}}\ MeanOfColumns(i)$;
  end
  return $predClass$;
end

4 Results and discussion

In this section, we present the results of our proposed anomaly detection procedures in terms of overall accuracies and analyze them with the related proper discussion. We use overall accuracy as a measure of how well our proposed anomaly detection methods identify and classify outages among four state categories. Accuracy metric in our study is simply the ratio of number of correct predictions to total number of predictions. We keep track of comparing the success of two deep learning architectures and the factors affecting their performance. Within this scope, we investigate the influences of various radius of interest values as 30, 40, and 50 meters, different shadowing standard deviations as 2, 5, and 8 dB as well as two different aggregation methods as ensemble averaging and majority voting.

A factor affecting four-category classification success emerges as the ROI which is related to the number of UEs involved in the analysis. In our analysis, we note a descending trend on accuracies along with increasing ROI values. Regardless of the aggregation method, we can see this trend in Fig. 4 which shows overall anomaly detection accuracies

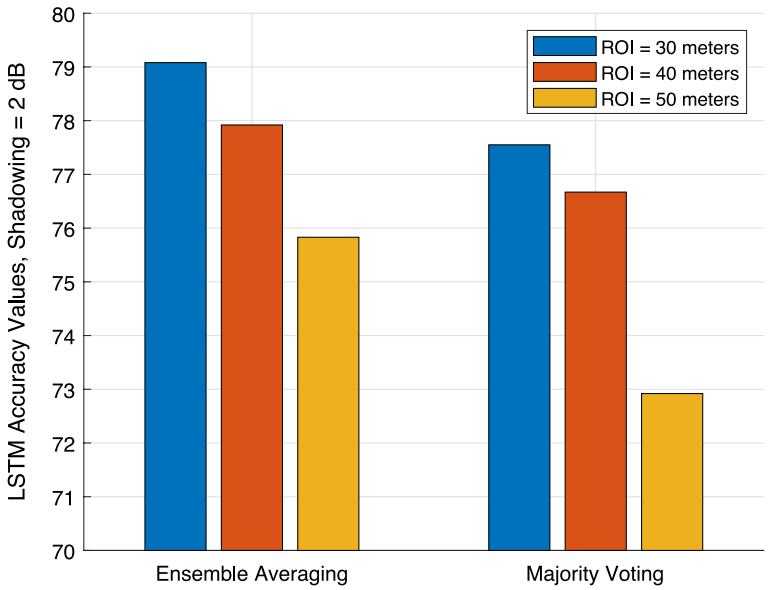

Fig. 4 LSTM accuracies at 2 dB shadowing

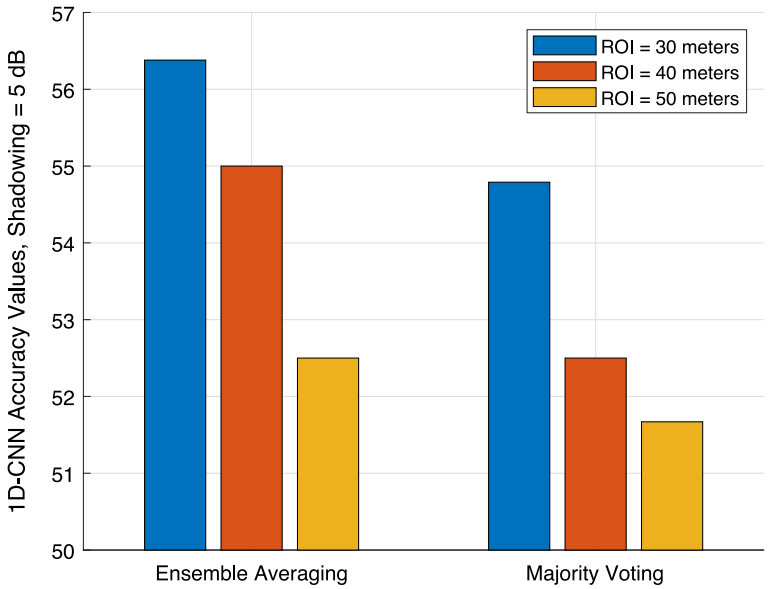

Fig. 5 1D CNN accuracies at 5 dB shadowing

of LSTM at 2 dB shadowing. Similarly, this can also be observed in Fig. 5 which shows the 5 dB shadowing accuracies of 1D CNN. In four-category classification scheme, it is worthy of noticing that as ROI gets bigger, although number of UEs subject to our classification and aggregation task gets more, the average information bearing quality per UE data gets less because the characteristics of data belonging to UEs placed near boundaries of ROI start to overlap with that of UEs belonging to macro BS due to path loss effects. Therefore, the bigger the radius we take, the noisier average UE data involved in our computations get. The results support this conclusion because in 50 meters of ROI,

although there are more UEs involved in classification, the accuracies all get lower and this holds for both aggregation decision methods.

In Fig. 6, we see LSTM and 1D CNN accuracies in the same plot for all ROI values under ensemble averaging. At 2 dB shadowing case, within 30 meters of ROI, 1D CNN (80.1%) outperforms LSTM (79.08%). On the other hand, LSTM gives accuracies as 59.18% and 50.83% outperforming 1D CNN with accuracies 56.38% and 47.43% for shadowing conditions at 5 dB and 8 dB, respectively. Here, we notice that 1D CNN is more successful at milder shadowing conditions whereas LSTM can better suit anomaly detection in harsher channel conditions. Special gates and cell memory structure of LSTM might stand more noise resilient in channel conditions with severe shadowing. However, in less shadowing within 30 meters, the patterns caused by anomalies on time sequenced UE data might become more salient under convolutional filtering power of 1D CNN. Under 40 and 50 meters of ROI, LSTM outperforms 1D CNN for all shadowing conditions except the accuracy within 50 meters for 5 dB in which they perform the same.

In addition, Fig. 6 also tells us that shadowing is another factor that affects classification and anomaly detection procedure. We notice that as shadow fading becomes harsher, under the same ROI, the classification accuracies all get less. This remarkable reduction obviously shows that harsher shadowing makes anomaly detection a harder task as the KPI data subject to classifiers start having more noisy nature.

Employed aggregation decision method, which is a necessary step for transforming UE probability distributions to FAP state prediction, also affects the overall anomaly detection performance. Anomaly classification accuracies within all ROI choices and both aggregation decision methods are shown in Fig. 7 for LSTM at 2 dB shadowing and in Fig. 8 for 1D CNN at 5 dB shadow fading. At all ROI choices, ensemble averaging performs better than majority voting for both deep learning architectures. This result

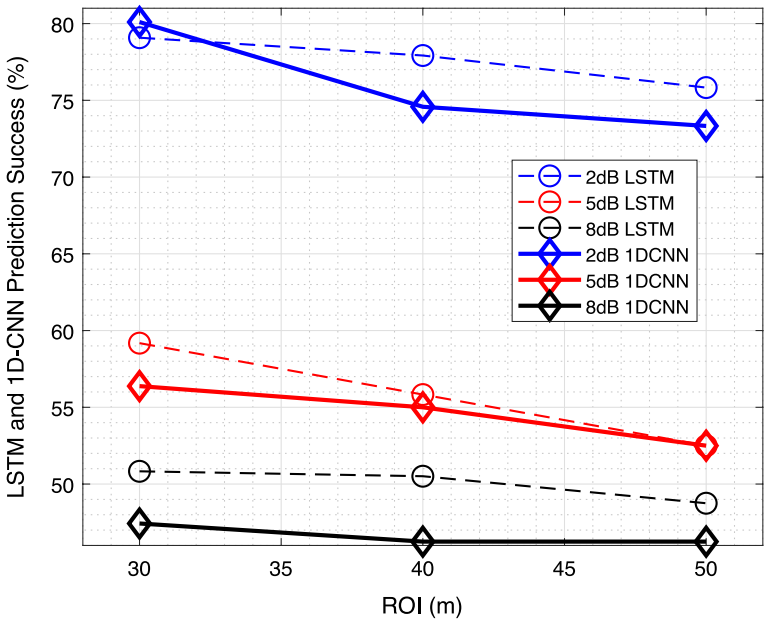

Fig. 6 Accuracies of LSTM and 1D CNN for all shadowing conditions and ROI choices

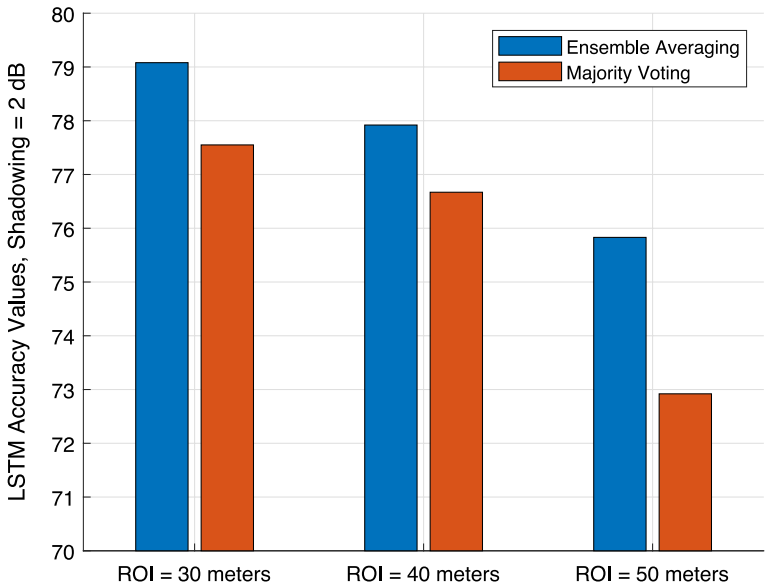

Fig. 7 Accuracies of LSTM at 2 dB shadowing

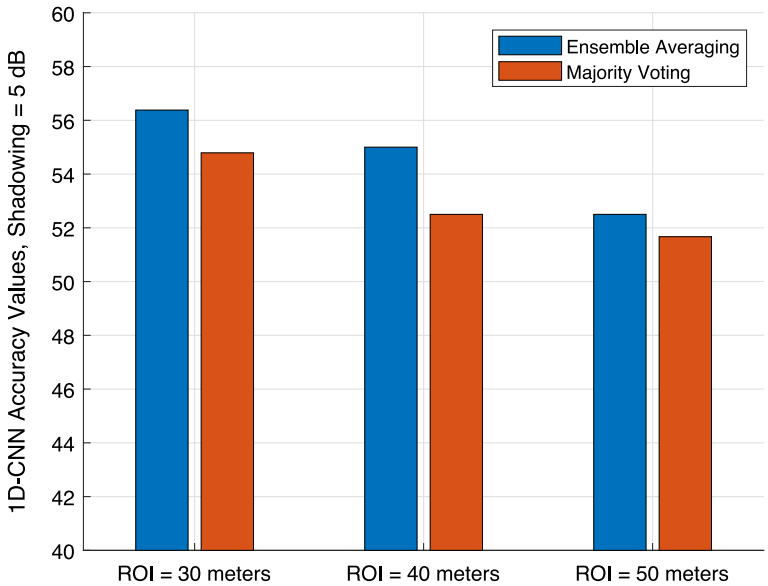

Fig. 8 Accuracies of 1D CNN at 5 dB shadowing

highlights the filtering power of ensemble averaging such that averaging the class posterior scores over the whole UE ensemble filters out potential errors of the model and also eliminates the outlying or faulty UE data, most probably corrupted by multi-path channel effects, so as to reduce misclassification [40].

In multi-label classification, recall rates for every category also account for the classifier. Recall rate is a label-based evaluation measure which gives the ratio of truly predicted samples of a category to total actual samples of that category as expressed in Equation (3) where TP indicates the true positives and FN indicates false negatives.

$$\text{Recall Rate} = \left[\frac{\text{TP}}{\text{TP} + \text{FN}}\right] \tag{3}$$

In Fig. 9, recall rates of every category are given for 30 meters of ROI under ensemble averaging. Here, we see again a different behavior between 2 dB shadowing and harsher shadowing conditions. At 2 dB shadowing, highest remarkable recall rates are achieved at healthy cases for both classifiers. On the other hand, at 5 dB and 8 dB shadowing conditions both classifiers give rather lower recall rates for degraded state. Therefore, our proposed four-category classifier extracts the catatonic and healthy states at the best, while it exhibits less success in identifying the degraded and crippled states that stand between healthy and catatonic states.

Besides focusing on four-category classification, the results can also be analyzed by classifying the FAPs into two categories. So, if we binarize four-category FAP classification structure into healthy and non-healthy states by merging degraded, crippled, and catatonic states into non-healthy state, the true positive rate (TPR) can also be another measure for assessment of our proposed two-state classification. TPR values simply indicate the rate of correctly identified positive non-healthy cases to all actual non-healthy cases. A good classifier is expected to have a TPR as high as possible.

In Fig. 10, TPR values of LSTM at 2 dB shadowing are shown for both aggregation algorithms and all ROI choices. For both aggregation methods, TPRs all increase along with increasing ROI which is the opposite of the case in four-category classification.

As shown in Fig. 11, for both aggregation methods, at 5 dB shadowing TPRs all decrease along with increasing ROI which is parallel to the case in four-category classification. Likewise, as shown in Fig. 13, there is also a decreasing trend of TPR values along with increasing ROI at 8 dB shadowing under ensemble averaging. Therefore, ROI affects the classification performance differently for two-state and four-state schemes at 2 dB shadowing and affects in the same way on other shadowing conditions. This different

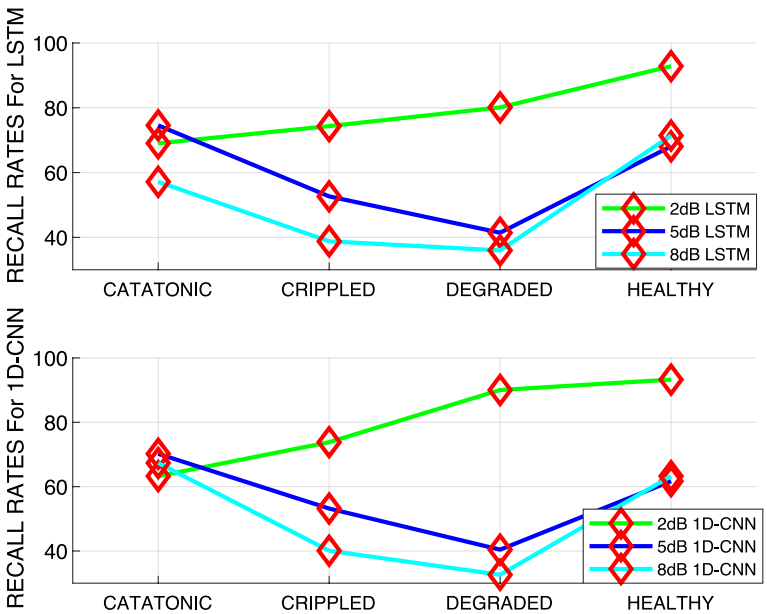

Fig. 9 Recall rates of LSTM and 1D CNN at 30 meters of ROI under ensemble averaging

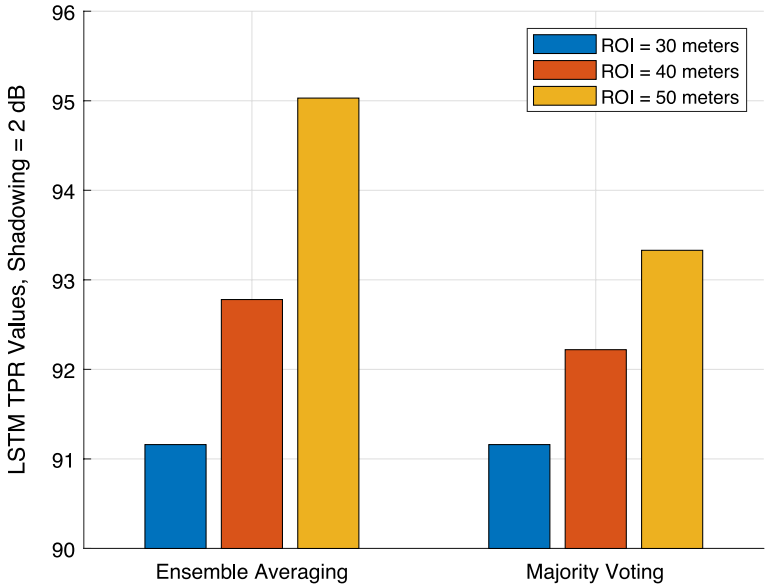

Fig. 10 TPRs of LSTM at 2 dB shadowing in two-state classification

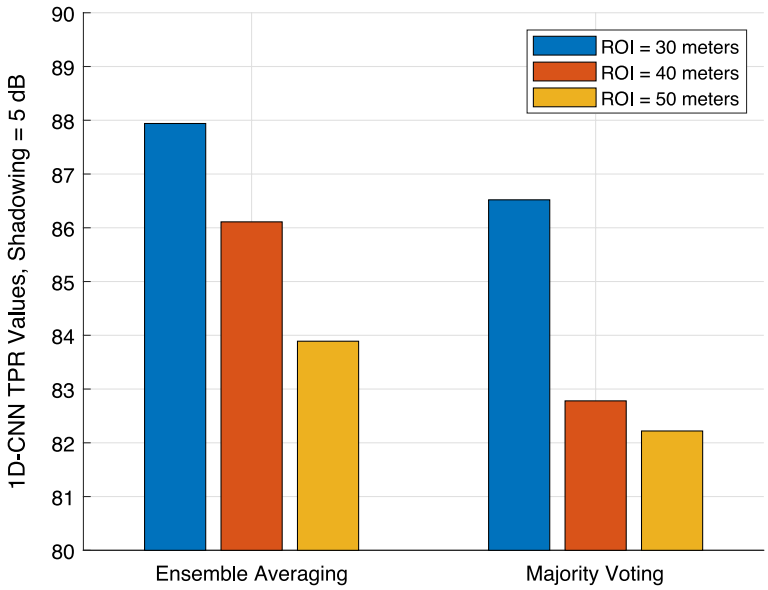

Fig. 11 TPRs of 1D CNN at 5 dB shadowing in two-state classification

behavior can be explained with the reason that in good signal conditions, every additional UE datum contains less noisy character. Thus, more qualified information leads to better detection of actual anomalous cases and prevents them from being erroneously classified as healthy. On the other hand, as shadowing increases to 5 and 8 dB, the information bearing quality of the UEs in the edges starts to decrease and severely overlaps with that of macro BS users. Classifiers start erroneous predictions on actual anomalous FAPs which lead to a reduction in TPRs. Therefore, misguidance of the classifiers due to higher shadowing yields lower TPRs since harsher channel conditions make outer UEs

less reliable for being incorporated in classification task for both classification schemes: two-state and four-category classification.

In two-state scheme, our proposed methods come up with very high TPRs when compared to four-category case accuracies since the internal misclassifications within the merged states all vanish. In our analysis, 1D CNN outperforms LSTM slightly in all cases for two-state scheme. As expected, the highest TPR is achieved with 2 dB shadowing condition with both classifiers having very close performances exceeding 95%. On the other hand, as shadow fading becomes harsher, TPR decreases. The best TPR is obtained within 50 meters of ROI at 2 dB shadowing, whereas the best TPR values are obtained within smaller ROI of 30 meters at harsher shadowing conditions with 5 and 8 dB for both deep learning architectures. This means that in finer channel conditions such as 2 dB shadowing, it becomes more advantageous for classifiers to utilize more samples by employing larger ROI.

In the literature, researchers study outage detection in cellular networks in various aspects. For example, in [26] researchers come up with similar TPR values in detection of two-state anomalies in a cellular network of macro BSs. Proposed methods in our study stand promising since we come up with high TPR values in the existence of extra difficulties such as user sparsity and vertical handovers posed by femtocells [6]. Likewise, in four-category FAP anomaly classification we reach accuracies more than 80% on the average which is comparable to the average results reached in [5] where more easily detectable macro BSs are also involved in. Moreover, our study does not require any data regarding neighboring cells as in [11] and does not require any KPI data preprocessing as in [28] and thus has a potential on being operated in run time applications for relatively reduced complexity.

For better understanding the performance of our proposed LSTM and 1D CNN schemes, using all the same UE sequences as input, we also employ a feed-forward neural network (FFNN) as a baseline, which is also used in outage detection [41]. The four-category classification accuracy values of LSTM, 1D CNN, and FFNN are shown in Fig. 12 under ensemble averaging aggregation. Proposed LSTM and 1D CNN methods both outperform existing FFNN approach in detecting and classifying anomalous states of ultra-dense femtocell networks in all shadowing conditions and for all ROI choices. By virtue of the gates and memory parts as well as cascaded layers, proposed deep learning methods may bring along computational complexity. However, proposed LSTM and 1D CNN methods have a much higher classification performance compared to the FFNN as used in [41] at the expense of an increase in computational cost.

We also emphasize the importance of aggregation decision methods employed in this study. For this purpose, we can consider on the PD_{UE} values which are the raw probability distribution vectors computed by the classifiers for each UE. So if we focus on the assigned class of UEs individually we can measure how much the applied aggregation method improves the decision about the state prediction of the FAPs of interest. Within this scope, in Table 2, we give prediction success of UE-based raw classification and that of FAPs following the aggregation decision process. We see that, in all cases, aggregation methods improve the prediction performance based on raw UE state probabilities.

By considering the results of this study, we can deduce that a smart way of anomaly management in ultra-dense FAP networks should be monitoring the signal quality

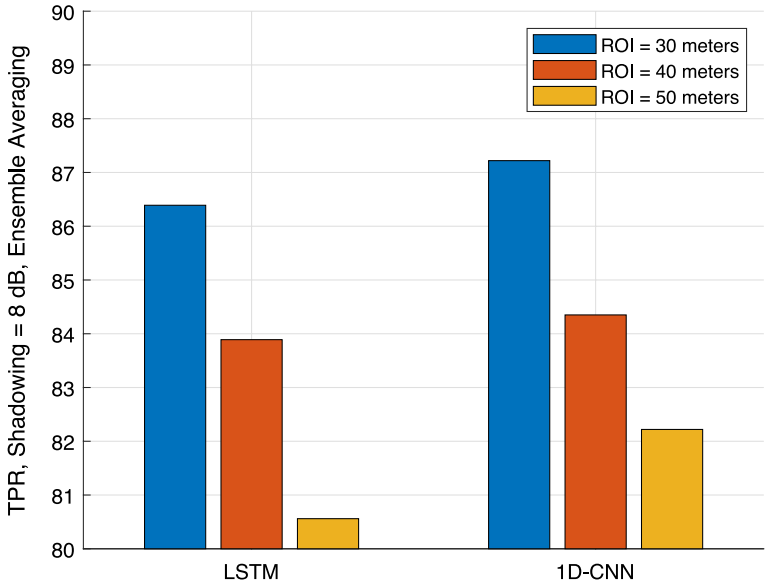

Fig. 12 Accuracies of proposed LSTM, proposed 1D CNN, and existing FFNN under ensemble averaging

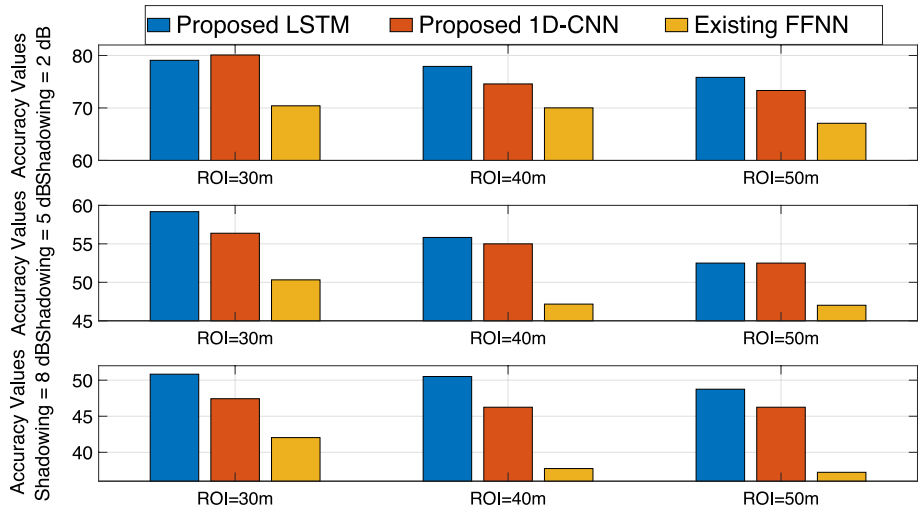

Fig. 13 TPRs of 1D CNN and LSTM at 8 dB shadowing with ensemble averaging in two-state classification

and channel conditions in radio access network and choose the appropriate type of the proposed deep learning method and optimum ROI accordingly. In that regard, in four-category classification, it is better to use LSTM in harsher channel conditions and use 1D CNN in good signal conditions. And in two-state scheme, employing larger ROI gives better accuracies in good signal conditions, whereas smaller radius is more suitable in harsher conditions. Therefore, OSS should keep track of channel conditions and determine the most appropriate way of detection method and suitable ROI for anomaly detection in cellular networks.

Our proposed framework mainly focuses on LSTM and 1D CNN. In addition to this, in the future our framework can also be well extended to hybrid schemes like

Table 2 Four-category accuracies: UE-based raw classification vs aggregation decision

	Accuracies (%)					
	2 dB		5 dB		8 dB	
	UE-based classification	Ensemble averaging	UE-based classification	Ensemble averaging	UE-based classification	Ensemble averaging
LSTM						
30m	65.75	79.08	51.40	59.18	43.50	50.83
40m	60.43	77.92	44.17	55.83	40.17	50.51
50m	55.70	75.83	41.04	52.50	39.83	48.75
1D CNN						
30m	65.75	80.10	48.60	56.38	42.13	46.43
40m	59.83	74.58	43.45	55.00	40.28	46.25
50m	54.77	73.33	41.67	52.50	36.54	46.25

RCNN structures. Moreover in the future planning, for the purpose of anomaly detection in cellular networks of femtocells, proposed time sequence monitoring and labeling tasks can also be handled together with other schemes. For example, hierarchical temporal memory (HTM) method can be a good candidate for its ability of being applicable to high-dimensional time sequences as covered in [42]. On the other hand, auto-encoder-based anomaly detection methods can also be well adopted to identification of cell outages in ultra-dense cellular networks [22, 43, 44].

5 Conclusion

In this study, we work on detecting the anomalies in femtocells and classifying them into four categories as healthy, degraded, crippled, and catatonic. By using KPI data of UEs along with LSTM and 1D CNN, we propose two different deep learning-based frameworks for efficient detection of anomalous FAP states which is much harder than detecting macro BS anomalies. We also showed the usage of aggregate decision approaches on UE prediction scores for further improving prediction results of FAP states. Our proposed method is capable of detecting anomalous cell states with more than 95% true positive rate and has an overall accuracy of more than 80% for four categorical state prediction of FAPs at the best. Having larger training datasets, working with higher user density and more sophisticated aggregate decision methods may help further increase classification accuracy so as to offer better insights into the anomaly detection tasks for future networks.

Abbreviations

LSTM	Long short-term memory
1D CNN	One-dimensional convolutional neural network
OSS	Operations support system
SON	Self-organizing networks
SINR	Signal-to-interference plus noise ratio
CQI	Channel quality indicator
FAP	Femto access point
3GPP	Third-Generation Partnership Project
COD	Cell outage detection
UE	User equipment
KPI	Key performance indicators
RNN	Recurrent neural network
RAN	Radio access network

TPR	True positive rate
FN	False negative
ROI	Femto access point radius of interest

Author contributions
HTO and AK contributed to the conception of the study, performed the simulations and data analyses, and wrote the manuscript. Both authors read and approved the final manuscript.

Availability of data and materials
The data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 3 February 2021 Accepted: 26 September 2022
Published online: 04 October 2022

References

1. ETSI : universal mobile telecommunications system (UMTS); LTE; telecommunication management; self-organizing networks, (SON) policy network resource model (NRM), integration reference point (IRP); information service (IS) (3GPP TS 28.628 version 11.0.0 release 11) (2013). ETSI, intelligent transport systems (ITS): vehicular communications; Part 6: Internet Integration; sub-part 1: transmission of IPv6 packets over GeoNetworking protocols (2011), Accessed 25 Jan 2021
2. J.M. DeAlmeida, C.F.T. Pontes, L.A. DaSilva, C.B. Both, J.J.C. Gondim, C.G. Ralha, M.A. Marotta, Abnormal behavior detection based on traffic pattern categorization in mobile networks. IEEE Trans. Netw. Serv. Manag. **18**(4), 4213–4224 (2021). https://doi.org/10.1109/TNSM.2021.3125019
3. F. Xie, D. Wei, Z. Wang, Traffic analysis for 5G network slice based on machine learning. EURASIP J. Wirel. Commun. Netw. **2021**(1), 1–15 (2021). https://doi.org/10.1186/s13638-021-01991-7
4. H.D. Trinh, E. Zeydan, L. Giupponi, P. Dini, Detecting mobile traffic anomalies through physical control channel fingerprinting: a deep semi-supervised approach. IEEE Access **7**, 152187–152201 (2019). https://doi.org/10.1109/ACCESS.2019.2947742
5. M. Alias, N. Saxena, A. Roy, Efficient cell outage detection in 5G hetnets using hidden Markov model. IEEE Commun. Lett. **20**(3), 562–565 (2016). https://doi.org/10.1109/LCOMM.2016.2517070
6. W. Wang, Q. Liao, Q. Zhang, COD: a cooperative cell outage detection architecture for self-organizing femtocell networks. IEEE Trans. Wirel. Commun. **13**(11), 6007–6014 (2014). https://doi.org/10.1109/TWC.2014.2360865
7. O. Onireti, A. Zoha, J. Moysen, A. Imran, L. Giupponi, M. Ali Imran, A. Abu-Dayya, A cell outage management framework for dense heterogeneous networks. IEEE Trans. Veh. Technol. **65**(4), 2097–2113 (2016). https://doi.org/10.1109/TVT.2015.2431371
8. O.G. Aliu, A. Imran, M.A. Imran, B. Evans, A survey of self organisation in future cellular networks. IEEE Commun. Surv. Tutor. **15**(1), 336–361 (2013). https://doi.org/10.1109/SURV.2012.021312.00116
9. P.V. Klaine, M.A. Imran, O. Onireti, R.D. Souza, A survey of machine learning techniques applied to self-organizing cellular networks. IEEE Commun. Surv. Tutor. **19**(4), 2392–2431 (2017). https://doi.org/10.1109/COMST.2017.2727878
10. I. de-la-Bandera, R. Barco, P. Muñoz, I. Serrano, Cell outage detection based on handover statistics. IEEE Commun. Lett. **19**(7), 1189–1192 (2015). https://doi.org/10.1109/LCOMM.2015.2426187
11. C.M. Mueller, M. Kaschub, C. Blankenhorn, S. Wanke, A Cell Outage Detection Algorithm Using Neighbor Cell List Reports, in *Self-organizing systems*. ed. by K.A. Hummel, J.P.G. Sterbenz (Springer, Berlin, Heidelberg, 2008), pp.218–229. https://doi.org/10.1007/978-3-540-92157-8_19
12. Q. Liao, M. Wiczanowski, S. Stańczak, Toward cell outage detection with composite hypothesis testing. in *2012 IEEE international conference on communications (ICC)*, pp. 4883–4887 (2012). https://doi.org/10.1109/ICC.2012.6364384
13. Y. Ma, M. Peng, W. Xue, X. Ji, A dynamic affinity propagation clustering algorithm for cell outage detection in self-healing networks. in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2266–2270 (2013). https://doi.org/10.1109/WCNC.2013.6554913
14. R.M. Khanafer, B. Solana, J. Triola, R. Barco, L. Moltsen, Z. Altman, P. Lazaro, Automated diagnosis for UMTS networks using bayesian network approach. IEEE Trans. Veh. Technol. **57**(4), 2451–2461 (2008). https://doi.org/10.1109/TVT.2007.912610
15. L. Bodrog, M. Kajo, S. Kocsis, B. Schultz, A robust algorithm for anomaly detection in mobile networks. in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–6 (2016). https://doi.org/10.1109/PIMRC.2016.7794573
16. P. Munoz, R. Barco, I. Serrano, A. Gomez-Andrades, Correlation-based time-series analysis for cell degradation detection in son. IEEE Commun. Lett. **20**(2), 396–399 (2016). https://doi.org/10.1109/LCOMM.2016.2516004
17. A. Gomez-Andrades, R. Barco, P. Muñoz, I. Serrano, Data analytics for diagnosing the RF condition in self-organizing networks. IEEE Trans. Mobile Comput. **16**(6), 1587–1600 (2017). https://doi.org/10.1109/TMC.2016.2601919
18. P. Munoz, R. Barco, E. Cruz, A. Gomez-Andrades, E.J. Khatib, N. Faour, A method for identifying faulty cells using a classification tree-based ue diagnosis in lte. EURASIP J. Wirel. Commun. Netw. **2017**(1), 1–20 (2017). https://doi.org/10.1186/s13638-017-0914-3

19. L. Fernandez Maimo, A.L. Perales Gomez, F.J. Garcia Clemente, M. Gil Perez, G. Martinez Perez, A self-adaptive deep learning-based system for anomaly detection in 5G networks. IEEE Access **6**, 7700–7712 (2018). https://doi.org/10.1109/ACCESS.2018.2803446
20. Y. Zuo, Y. Wu, G. Min, C. Huang, K. Pei, An intelligent anomaly detection scheme for micro-services architectures with temporal and spatial data analysis. IEEE Trans. Cogn. Commun. Netw. **6**(2), 548–561 (2020). https://doi.org/10.1109/TCCN.2020.2966615
21. A. Bhardwaj, F. Al-Turjman, V. Sapra, M. Kumar, T. Stephan, Privacy-aware detection framework to mitigate new-age phishing attacks. Comput. Electr. Eng. **96**, 107546 (2021). https://doi.org/10.1016/j.compeleceng.2021.107546
22. M. Savic, M. Lukic, D. Danilovic, Z. Bodroski, D. Bajović, I. Mezei, D. Vukobratovic, S. Skrbic, D. Jakovetić, Deep learning anomaly detection for cellular IoT with applications in smart logistics. IEEE Access **9**, 59406–59419 (2021). https://doi.org/10.1109/ACCESS.2021.3072916
23. A. Asghar, H. Farooq, H.N. Qureshi, A. Abu-Dayya, A. Imran, Entropy field decomposition based outage detection for ultra-dense networks. IEEE Access, 1 (2021). https://doi.org/10.1109/ACCESS.2021.3056551
24. J.M. DeAlmeida, C.F.T. Pontes, L.A. DaSilva, C.B. Both, J.J.C. Gondim, C.G. Ralha, M.A. Marotta, Abnormal behavior detection based on traffic pattern categorization in mobile networks. IEEE Trans. Netw. Serv. Manag. **18**(4), 4213–4224 (2021). https://doi.org/10.1109/TNSM.2021.3125019
25. D. Mulvey, C.H. Foh, M. Ali Imran, R. Tafazolli, Cell coverage degradation detection using deep learning techniques. in *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 441–447 (2018). https://doi.org/10.1109/ICTC.2018.8539449
26. B. Hussain, Q. Du, A. Imran, M.A. Imran, Artificial intelligence-powered mobile edge computing-based anomaly detection in cellular networks. IEEE Trans. Ind. Inf. **16**(8), 4986–4996 (2020). https://doi.org/10.1109/TII.2019.2953201
27. J.C. Ikuno, M. Wrulich, M. Rupp, System level simulation of LTE networks. in *2010 IEEE 71st Vehicular Technology Conference*, pp. 1–5 (2010). https://doi.org/10.1109/VETECS.2010.5494007
28. S.M.A. Al Mamun, J. Valimaki, Anomaly detection and classification in cellular networks using automatic labeling technique for applying supervised learning. Proc. Comput. Sci. **140**, 186–195 (2018). https://doi.org/10.1016/j.procs.2018.10.328
29. S. Hochreiter, J. Schmidhuber, Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997). https://doi.org/10.1162/neco.1997.9.8.1735
30. A. Aggarwal, A. Rani, P. Sharma, M. Kumar, A. Shankar, M. Alazab, Prediction of landsliding using univariate forecasting models. Int. Technol. Lett. **5**(1), 209 (2022). https://doi.org/10.1002/itl2.209
31. X. Zhang, Y. Zheng, W. Liu, Z. Wang, A hyperspectral image classification algorithm based on atrous convolution. EURASIP J. Wirel. Commun. Netw. (2019). https://doi.org/10.1186/s13638-019-1594-y
32. A. Thakkar, R. Lohiya, Analyzing fusion of regularization techniques in the deep learning-based intrusion detection system. Int. J. Intell. Syst. **36**(12), 7340–7388 (2021). https://doi.org/10.1002/int.22590
33. C. Wu, W. Li, Enhancing intrusion detection with feature selection and neural network. Int. J. Intell. Syst. **36**(7), 3087–3105 (2021). https://doi.org/10.1002/int.22397
34. I.J. Goodfellow, Y. Bengio, A. Courville, Deep Learning. (MIT Press, Cambridge, MA, 2016). http://www.deeplearningbook.org
35. A.A. Suárez León, J.R. Núñez Alvarez, 1D convolutional neural network for detecting ventricular heartbeats. IEEE Latin Am. Trans. **17**(12), 1970–1977 (2019). https://doi.org/10.1109/TLA.2019.9011541
36. T. Mir, L. Dai, Y. Yang, W. Shen, B. Wang, Optimal femtocell density for maximizing throughput in 5G heterogeneous networks under outage constraints. in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–5 (2017). https://doi.org/10.1109/VTCFall.2017.8288059
37. S.A. Mahmud, G.M. Khan, M. Zafar, K. Ahmad, N. Behttani, A survey on femtocells: benefits deployment models and proposed solutions. J. Appl. Res. Technol. **11**, 733–754 (2013). https://doi.org/10.1016/S1665-6423(13)71582-7
38. P. Chithaluru, A.-T. Fadi, M. Kumar, T. Stephan, Mtcee-lln: multi-layer threshold cluster-based energy efficient low power and lossy networks for industrial internet of things. IEEE Int. Things J. (2021). https://doi.org/10.1109/JIOT.2021.3107538
39. P.K. Chithaluru, M.S. Khan, M. Kumar, T. Stephan, Eth-leach: an energy enhanced threshold routing protocol for WSNs. Int. J. Commun. Syst. **34**(12), 4881 (2021). https://doi.org/10.1002/dac.4881
40. H. Zheng, Y. Cheng, H. Li, Investigation of model ensemble for fine-grained air quality prediction. China Commun. **17**(7), 207–223 (2020). https://doi.org/10.23919/J.CC.2020.07.015
41. W. Feng, Y. Teng, Y. Man, M. Song, Cell outage detection based on improved BP neural network in LTE system. in *11th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2015)*, pp. 1–5 (2015). https://doi.org/10.1049/cp.2015.0710
42. C. Wang, Z. Zhao, L. Gong, L. Zhu, Z. Liu, X. Cheng, A distributed anomaly detection system for in-vehicle network using HTM. IEEE Access **6**, 9091–9098 (2018). https://doi.org/10.1109/ACCESS.2018.2799210
43. Z. Cheng, S. Wang, P. Zhang, S. Wang, X. Liu, E. Zhu, Improved autoencoder for unsupervised anomaly detection. Int. J. Intell. Syst. **36**(12), 7103–7125 (2021). https://doi.org/10.1002/int.22582
44. F. Zhang, H. Fan, R. Wang, Z. Li, T. Liang, Deep dual support vector data description for anomaly detection on attributed networks. Int. J. Intell. Syst. **37**(2), 1509–1528 (2022). https://doi.org/10.1002/int.22683

Publisher's Note