In: 2023 5th International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 20-23 February 2023, Bali, Indonesia
DOI: 10.1109/ICAIIC57133.2023.10067061

Federated Learning of Wireless Network Experience Anomalies using Consumer Sentiment

1st Weisi Guo
Centre for Auto. Cyberphysical Sys.
Cranfield University
Bedford, UK
weisi.guo@cranfield.ac.uk

2nd Bailu Jin
Centre for Auto. Cyberphysical Sys.
Cranfield University
Bedford, UK
bailu.jin@cranfield.ac.uk

3rd Schyler C. Sun
Centre for Auto. Cyberphysical Sys.
Cranfield University
Bedford, UK
schyler.sun@cranfield.ac.uk

4th Yue Wu
East China University of Science & Technology
Shanghai, China
yuewu@ecust.edu.cn

5th Weijie Qi
Ranplan
Cambridge, UK
weijie.qi@ranplanwireless.com

6th Jie Zhang
Ranplan
Cambridge, UK
jie.zhang@ranplanwireless.com

***Abstract*—In wireless networks, consumer experience is important for both short monitoring of the Quality of Experience (QoE) as well as long term customer retainment. Current 4G and 5G networks are not equipped to measure QoE in an automated way, and experience is still reported through traditional customer care and drive-testing. In recent years, large-scale social media analytics has enabled researchers to gather statistically significant data on consumer experience and correlate them to major events such as social celebrations or significant network outages. However, the translational pathway from languages to topic-specific emotions (e.g., sentiment) to detecting anomalies in QoE is challenging. This challenge lies in two issues: (1) the social experience data remains sparsely distributed across space, and (2) anomalies in experience jump across sub-topic spaces (e.g., from data rate to signal strength).**

Here, we solved these two challenges by examining the spectral space of experience across topics using federated learning (FL) to identify anomalies. This can inform telecom operators to pay attention to potential network demand or supply issues in real time using relatively sparse and distributed data. We use real social media data curated for our telecommunication projects across London and the United Kingdom to demonstrate our results. FL was able to achieve 74-92% QoE anomaly detection accuracy, with the benefit of 30-45% reduce data transfer and preserving privacy better than raw data transfer.

***Index Terms*—federated learning, wireless network, quality of experience, sentiment analysis, social media**

I. INTRODUCTION

The dominance of multimedia services across mobile devices (e.g., Twitter, Wechat, Whatsapp, Instagram, TikTok, Pokemon Go) means we have moved away from call and SMS text driven services to a more diverse user experience based services. According to International Telecommunication Union (ITU), the definition for Quality-of-Experience (QoE) is "the degree of delight or annoyance of the user of an application or service." QoE is more important than ever [1], but yet there very few ways to gauge QoE by the telecommunication operator in current practice (there are no 5G standards) for implementing QoE [2]. Whilst the app-base service themselves gauge general consumer experience of the platform, second-by-second network experience is important to understand both consumer demand dynamics and how the network supplies service-specific capacity. Whilst innovative proposals exist to use affective computing to monitor human emotional responses to services via the smartphone [3], integrating these services and getting consumer participation in a privacy preserving manner, remain out of reach for most operators and platform providers.

We would like to thank funding from EC H2020 DAWN4IoE - Data Aware Wireless Network for Internet-of-Everything (778305). We thank our other industrial and academic project partners WINGS-ICT and University of Iowa for their advice.

A. Related Work

Current 4G and 5G networks (as well as other wireless systems, Wi-Fi, Internet of Things/Everything) are not equipped to measure QoE in an automated way, and experience is still reported through traditional customer care and drive-testing. This is slow and passive, often accumulating a large number of negative reports before investigation. In propagation dominated networks, this practice is effective. However, in dynamic networks where spectrum is cognitively used to adapt to spike demand, this approach is no longer fit for purpose. Proactive network based on real-time data is needed for the consumers of today and tomorrow [4], [5]. The latter approach uses machine learning to directly infer QoE from diverse QoS data, but cannot capture the nature and context of the human experience.

In recent years, large-scale social media analytics has enabled researchers to gather statistically significant data on consumer experience and correlate them to major events such as social celebrations or significant network outages. For example, general perceptions of 5G and technology has been data mined from Twitter [15]. Detailed high fidelity sub-topic-specific research started in 2016 with correlating social media data with wireless network traffic data [6] to show the importance of leveraging on social media as a proxy data source.

© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works

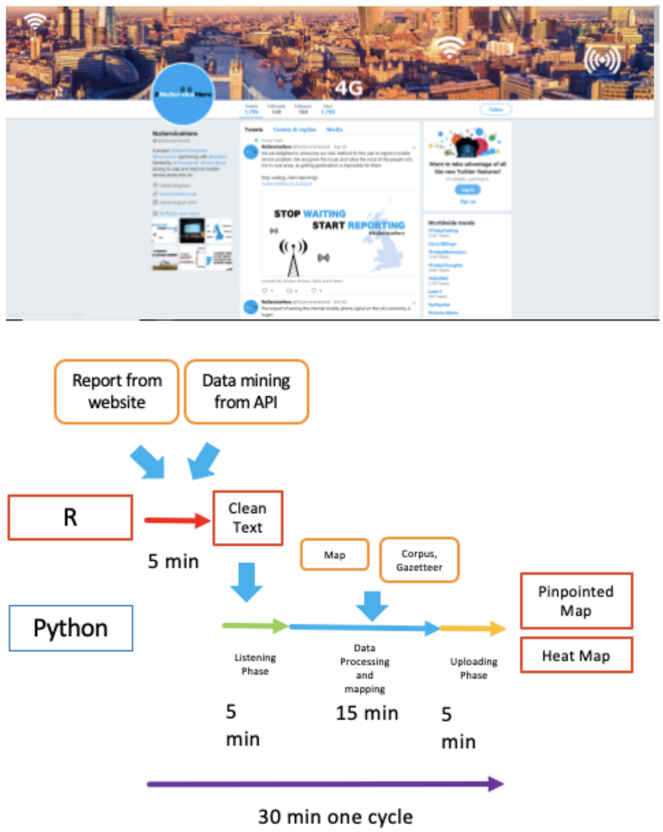

Fig. 1. (top) Social data acquisition and (bottom) sentiment analytics for QoE mapping.

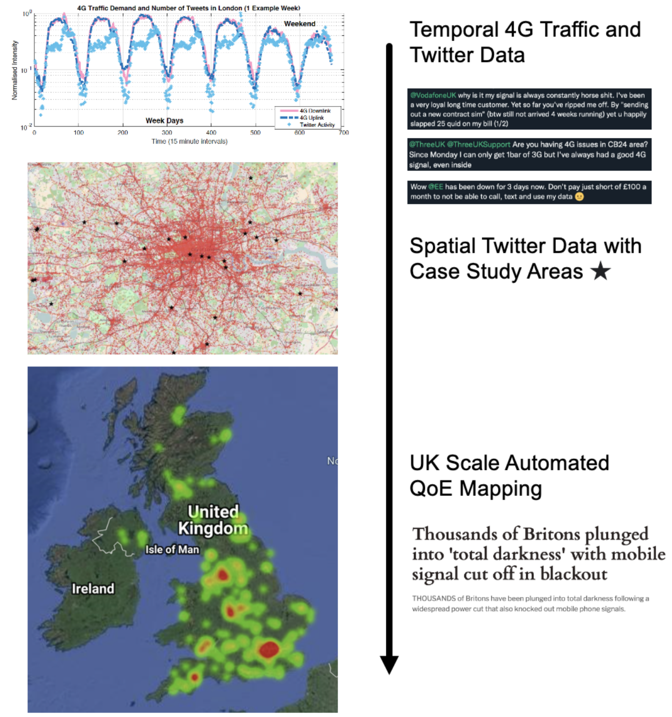

Fig. 2. (top) Temporal 4G Traffic and Twitter Data, (middle) Spatial Twitter Data with Case Study Areas, and (bottom) UK Scale Automated QoE Mapping.

However, the translational pathway from languages to topic-specific emotions (e.g., sentiment) to detecting anomalies in QoE is challenging. Our own work in 2017-19 examined how we can mine Twitter sentiment towards telecom sub-topics (e.g., signal reception, data rate) as a way of understanding real-time consumer QoE [7]. We expanded this work to paint contexts through geographic specific case studies, event-based case studies (e.g., O2 telecom blackout [8]). Nonetheless, we still identified two major challenges:

- the social QoE data remains sparsely distributed across space, and
- anomalies in QoE jump across sub-topic spaces (e.g., each person will discuss data rate to signal strength).

That means it is very hard to get a statistically significant understanding across a large service area (e.g., London) what the real problem is and how significant is it? Whilst we can aggregate all the social media data to a single source for processing, this will require significantly large QoE intelligence gathering program and raise privacy concerns. Many of the Tweets maybe private, or not have the environmental context (e.g., location) sub-field data.

B. Innovation

Here, we propose to solve these two challenges by examining the spectral space of experience across topics using the following innovations:

- Federated Learning (FL) to identify QoE anomalies: this has the advantage of local users or nodes can analyse local data in a privacy preserving way without sending raw QoE data. Certainly this has been examined in smart city contexts recently [9].
- Use the FL to examined topic spectral diagrams in order to get an understanding of the general QoE problem across sub-topic spaces.

Combined together, we can inform telecom operators to pay attention to potential network demand or supply issues in real time using relatively sparse and distributed data. We use real social media data curated for our telecommunication projects across London and the United Kingdom to demonstrate our results.

II. Social Data and Sentiment Analysis

A. Data Campaign

Mobile wireless connectivity black-spots and poor service outages is often hidden, under-reported, and lead to severe consumer dissatisfaction and digital economy problems. There is an urgent requirement for improving the market competitiveness of the UK commerce by providing social data to enhance the existing mobile services (e.g. 4G, 5G) and, to guide the and future 5G Advanced investment. In our previous work, we used 2 strategies to obtain both users' qualitative experience of mobile network and their geographical information:

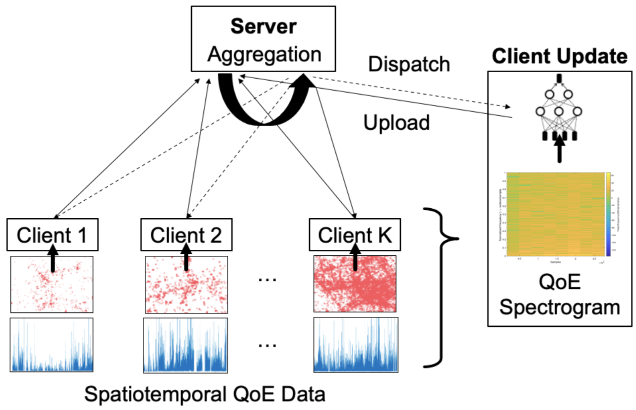

Fig. 3. Federated learning of QoE using local sparse data. Spectrogram of sub-topic sentiments through a CNN can detect anomalies in QoE.

- general large-scale data mining of QoE reporting on Twitter (validated through case studies using secondary cell signal data) [7], and
- encourage users to report QoE through dedicated social influence campaigns by @NoServiceHere Twitter account using the hashtag NoServiceHere so that we can identify specific circumstances and context more accurately [8] - see Fig. 1top.

These studies were a joint university and industry (Ranplan UK) initiative. A work flow diagram of the social media data gathering and the following sentiment analysis for QoE is given in Fig. 1bottom. We are able to gather social media data from both a purpose built website account and the Twitter API, use R to process the text data to clean out spelling mistakes and typos, and then use Python code to perform the machine learning pipeline discussed below. This then allows us to map and contextualise the QoE data.

B. Sentiment Analysis for QoE

In the sentiment analysis, we first identify the topic being discussed for two purposes: (1) to separate out similar key phrases but in different domains (e.g., "a bad reception" can mean wireless signal or wedding); and (2) to identify sub-topics within wireless (e.g., "SMS send failure", "Signal but No Service"). This allows us to create wireless relevant sub-topics to then focus on the sentiment towards them as well as the user context. From previous experience (see [6], [7]) we found that n-gram grammar models that capture double negatives and more sophisticated grammar than keyword based approaches ($n = 1$), required $n = 4$ for good accuracy.

For sentiment analysis on the sub-topics, we have created our own Corpus and Sentiment Analysis System [6], [7] and apply the following steps: (i) n-gram TDSA model combined with corpus filtering, (ii) machine learning classifier with 3 choices (Naive Bayes, Support Vector Machine, and a Recurrent Neural Network). Please see [6], [7] for our previous work details. We used a relatively simple individual user's sentiment scale of positive (+1) and negative (-1) to define our

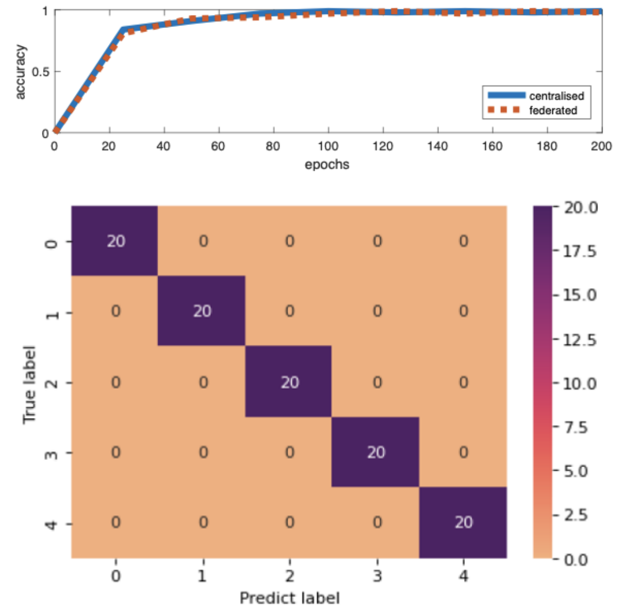

Fig. 4. Centralised learning (CL) and federated learning (FL) training and confusion matrix.

QoE for this study, and therefore track a discrete QoE variation across sub-topics. We then use a Short-Time Fourier Transform (STFT) to capture how spectral features may exhibit frequency anomalies in QoE rather than examine the time series directly (similar ideas have been used in sentiment analysis of speech for example [10], [13]).

C. Anomaly Detection in QoE

Our Convolution Neural Network (CNN) is trained on case study data sets that are commercially sensitive, whereby we have known network outage events or bad coverage case studies that allow us to train the CNN to identify the equivalent spatiotemporal QoE anomaly profiles. We then use the trained federated CNN to detect anomalies across a wider urban area. The CNN model used is similar to VGG16 ImageNet (https://keras.io/api/applications/vgg/).

III. Federated Learning of QoE Anomaly

In order to avoid aggregating raw consumer data across a large area (a city of several RANs and hundreds of base stations), we will use local learning (e.g., on each base station) to act as a client k. There will be K clients across a large city, where local learning of the QoE spectrograms across all sub-topics achieves privacy preserving methods to monitor a whole city. Federated learning is a distributed framework in which the communication network uploads and downloads learning weights and gradients rather than the raw data. In contrast to traditional distributed learning where the compute nodes and the central server are commonly in the same geographical location, federated learning clients are located in different geographical locations and do not have a dedicated or synchronised communication channel. Depending on the

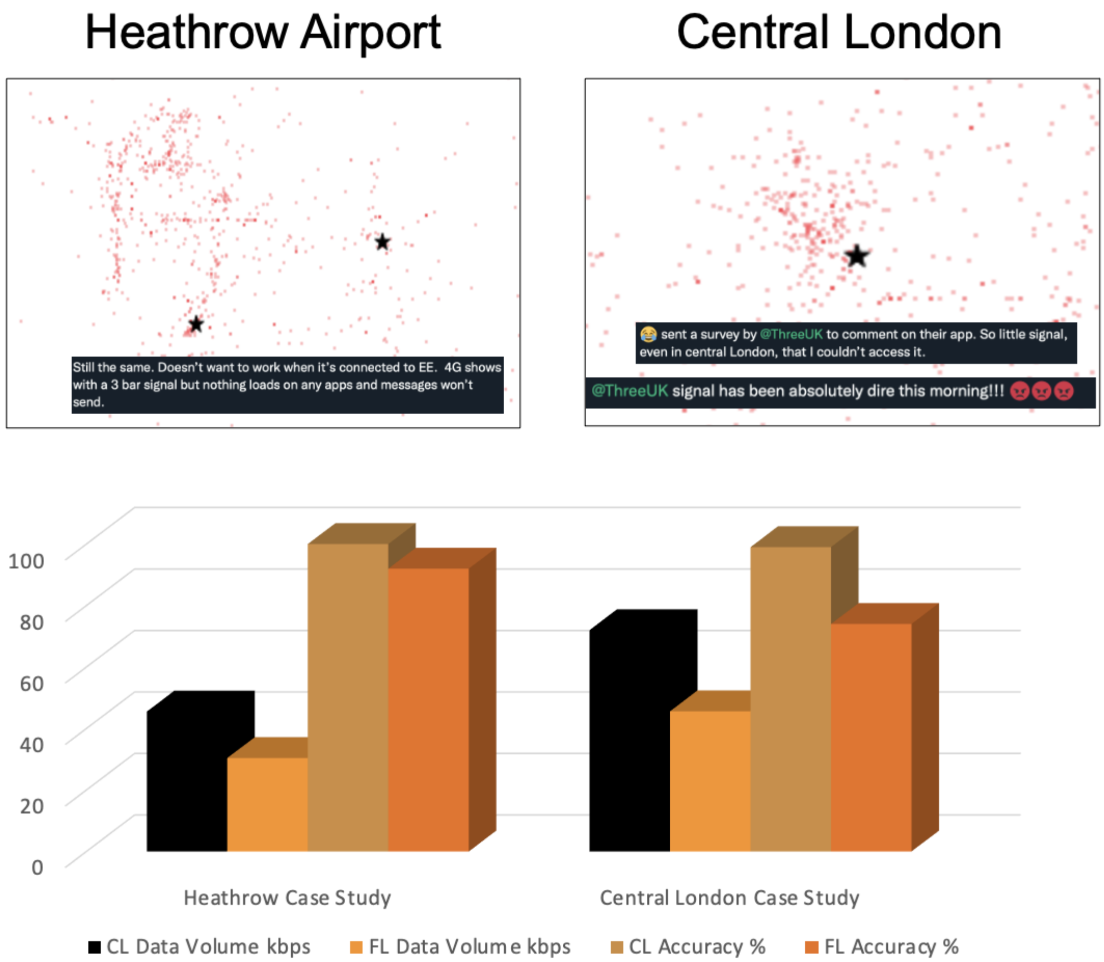

Fig. 5. Federated learning case studies and comparison results. Each case is approximately 50 QoE data points.

distribution of the data, federated learning can be divided into three categories, horizontal federated learning, vertical federated learning and federated transfer learning.

Here we implement horizontal federation learning (FL) [12], which has 3 steps: (1) Dispatch: The client downloads the model from the central server, (2) Aggregation: Each client trains the model using local data, encrypts the gradients and uploads them to the server, which aggregates the gradients from each user to update the model parameters, and (3) Local update: Each client updates its own model. We employ the Federated Averaging (FedAvg) is an algorithm whose main idea is to achieve a reduction in communication consumption by increasing the number of client-side computations [11]. In the FedAvg process, the central server will randomly select $j \in K$ clients to sample, and the gradient update of these clients will be averaged as the gradient of the global model of the server, and finally the server will issue the gradient of the global model to the client model and carry out a continuous loop. There are four steps in FedAvg:

1) At step t of each iteration round, the server sends the current global model parameters m_k to the client
2) Unselected clients update via Adam optimiser according to m_k
3) Selected clients update local parameters $m_k(t+1)$
4) At iteration step $t+1$, the server calculates the weighted average $\overline{m}_k(t+1)$ and gets the new update to the server: $\sum_{k=1}^{K} \frac{n_k}{n} m_k(t+1)$

Fig. 4 shows the federated learning (FL) vs. centralised learning (CL) training accuracy over 20 example case studies and the confusion matrix.

IV. Results & Discussions

We perform large scale surveillance using distributed FL clients based on base station locations ($K = 96$ location areas), and use anomaly detection to identify 2 case studies. Each case is approximately 50 QoE data points. The first case study contains two closely located QoE anomalies at the Terminal 5 and Terminal 1-3 of Heathrow Airport, showing persistent dissatisfaction with wireless coverage, due to spectrum “congestion and over demand“ (sub-topic). The second case study centres around Waterloo training station and contains both “congestion and over demand“ (sub-topic) during train cancellations, as well as persistent “poor signal strength“ (sub-topic) inside the station. The CL approach was able to achieve 99-100% accuracy in these test cases when compared to validation data we have OpenSignal. FL here, has significantly lower accuracy 74-92%, but offers the benefit of reduced data flow volume from clients to server. Compared to CL, we see roughly 30-45% reduction in data transfer as gradients and parameters have lower data requirements than raw QoE social media data.

In Table I, we offer a full comparison of different CL vs. FL learning structures (all using VGG16 CNN for anomaly detection), and different NLP engines for the QoE classification. We show that RNN are significantly better and the improved data privacy and data reduction tradeoff for FL is strong.

TABLE I
PERFORMANCE COMPARISON OF CL AND FL IN DIFFERENT QOE ANOMALY DETECTION CASES. EACH CASE IS APPROXIMATELY 50 QOE DATA POINTS.

Learning Structure	NLP Engine	Case Study Area	Accuracy	Data Uplink
CL with VGG16 CNN	Naive Bayes	Heathrow Airport	74%	45.5 kbps
CL with VGG16 CNN	SVM	Heathrow Airport	86%	45.5 kbps
CL with VGG16 CNN	RNN	Heathrow Airport	100%	45.5 kbps
FL with VGG16 CNN	RNN	Heathrow Airport	92%	30.4 kbps
CL with VGG16 CNN	Naive Bayes	Central London	73%	72 kbps
CL with VGG16 CNN	SVM	Central London	81%	72 kbps
CL with VGG16 CNN	RNN	Central London	99%	72 kbps
FL with VGG16 CNN	RNN	Central London	74%	45.6 kbps

V. CONCLUSIONS & FUTURE WORK

In wireless networks, consumer experience is important for both short monitoring of the Quality of Experience (QoE) as well as long term customer retainment. Current 4G and 5G networks are not equipped to measure QoE in an automated way, and experience is still reported through traditional customer care and drive-testing. In recent years, large-scale social media analytics has enabled researchers to gather statistically significant data on consumer experience and correlate them to major events such as social celebrations or significant network outages. However, the translational pathway from languages to topic-specific emotions (e.g., sentiment) to detecting anomalies in QoE is challenging. This challenge lies in two issues: (1) the social experience data remains sparsely distributed across space, and (2) anomalies in experience jump across sub-topic spaces (e.g., from data rate to signal strength).

Here, we solved these two challenges by examining the spectral space of experience across topics using federated learning (FL) to identify anomalies. This can inform telecom operators to pay attention to potential network demand or supply issues in real time using relatively sparse and distributed data. We use real social media data curated for our telecommunication projects across London and the United Kingdom to demonstrate our results in 2 case studies where we were able to detect significant QoE anomalies. In the case studies we examined, FL was able to achieve 74-92% QoE anomaly detection accuracy, with the benefit of 30-45% reduce data transfer and preserving privacy better than raw data transfer.

In future work, we will examine how we can obtain higher dimensional definitions of QoE beyond +1 and -1 unidimensional sentiments per person. We may wish to consider how we can be application or service slice specific, and how we can integrate other affective computing methods with social media analytics [14], and how this can drive network management [16].

REFERENCES

[1] B. Agarwal, M. A. Togou, M. Ruffini and G. -M. Muntean, “QoE-Driven Optimization in 5G O-RAN Enabled HetNets for Enhanced Video Service Quality,“ in IEEE Communications Magazine, 2022.

[2] R. D. Mardian, M. Suryanegara and K. Ramli, “Measuring Quality of Service (QoS) and Quality of Experience (QoE) on 5G Technology: A Review,“ 2019 IEEE International Conference on Innovative Research and Development (ICIRD), 2019

[3] B. Dudin, N. A. Ali, A. Radwan and A. -E. M. Taha, “Resource Allocation with Automated QoE Assessment in 5G/B5G Wireless Systems,“ in IEEE Network, vol. 33, no. 4, pp. 76-81, July/August 2019.

[4] B. Ma, W. Guo and J. Zhang, “A Survey of Online Data-Driven Proactive 5G Network Optimisation Using Machine Learning,“ in IEEE Access, vol. 8, pp. 35606-35637, 2020.

[5] U. S. Hashmi et al., “Towards Real-Time User QoE Assessment via Machine Learning on LTE Network Data,“ 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), 2019.

[6] B. Yang, W. Guo, B. Chen, G. Yang and J. Zhang, “Estimating Mobile Traffic Demand Using Twitter,“ in IEEE Wireless Communications Letters, vol. 5, no. 4, pp. 380-383, Aug. 2016.

[7] W. Qi, R. Procter, J. Zhang and W. Guo, “Mapping Consumer Sentiment Toward Wireless Services Using Geospatial Twitter Data,“ in IEEE Access, vol. 7, pp. 113726-113739, 2019.

[8] W. Qi, W. Guo, R. Procter and J. Zhang, “Geo-Tagging Quality-of-Experience Self-Reporting on Twitter to Mobile Network Outage Events,“ 2019 IEEE International Smart Cities Conference (ISC2), 2019

[9] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li and H. Vincent Poor, “Federated Learning for Internet of Things: A Comprehensive Survey,“ in IEEE Communications Surveys & Tutorials, vol. 23, no. 3, pp. 1622-1658, thirdquarter 2021.

[10] A. M. Badshah, J. Ahmad, N. Rahim and S. W. Baik, “Speech Emotion Recognition from Spectrograms with Deep Convolutional Neural Network,“ 2017 International Conference on Platform Technology and Service (PlatCon), 2017

[11] X. Li, K. Huang, W. Yang, S. Wang, Z. Zhang, “On the Convergence of FedAvg on Non-IID Data,“ 2020 International Conference on Learning Representations, 2020.

[12] P. Chhikara, P. Singh, R. Tekchandani, N. Kumar and M. Guizani, “Federated Learning Meets Human Emotions: A Decentralized Framework for Human?Computer Interaction for IoT Applications,“ in IEEE Internet of Things Journal, vol. 8, no. 8, pp. 6949-6962, 15 April15, 2021

[13] S. Amiriparian, N. Cummins, S. Ottl, M. Gerczuk and B. Schuller, “Sentiment analysis using image-based deep spectrum features,“ 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), 2017

[14] M. Chen, Y. Zhang, Y. Li, S. Mao and V. C. M. Leung, “EMC: Emotion-aware mobile cloud computing in 5G,“ in IEEE Network, vol. 29, no. 2, pp. 32-38, March-April 2015

[15] I. K. A. Aryadinata, D. Pangesti, G. B. Anugerah, I. E. Aditya and Y. Ruldeviyani, ”Sentiment Analysis of 5G Network Implementation in Indonesia Using Twitter Data,” 2021 6th International Workshop on Big Data and Information Security (IWBIS), 2021

[16] A. A. Barakabitze et al., “QoE Management of Multimedia Streaming Services in Future Networks: A Tutorial and Survey,“ in IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 526-565, Firstquarter 2020

Cranfield University

CERES https://dspace.lib.cranfield.ac.uk

School of Aerospace, Transport and Manufacturing (SATM) Staff publications (SATM)

2023-03-23

Federated learning of wireless network experience anomalies using consumer sentiment

Guo, Weisi

IEEE

Guo W, Jin B, Sun SC, et al., (2023) Federated learning of wireless network experience anomalies using consumer sentiment. In: 2023 5th International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 20-23 February 2023, Bali, Indonesia

https://doi.org/10.1109/ICAIIC57133.2023.10067061

Downloaded from Cranfield Library Services E-Repository