ANALYSIS AND MITIGATION OF VOCAL EFFORT VARIATIONS IN SPEAKER RECOGNITION

Mahesh Kumar Nandwana[1], *Mitchell McLaren*[1], *Luciana Ferrer*[2], *Diego Castan*[1], *Aaron Lawson*[1]

[1]Speech Technology and Research Laboratory, SRI International, Menlo Park, California, USA
[2]Instituto de Investigacón en Ciencias de la Computación, UBA-CONICET, Argentina

ABSTRACT

In this work, we assess the impact of vocal effort on discrimination and calibration performance of a state-of-the-art speaker recognition system. We analyze three levels of vocal effort (low, normal, and high) from the SRI-FRTIV corpus. We use a deep neural network (DNN) speaker embeddings system with probabilistic linear discriminant analysis (PLDA) and find that vocal effort variation significantly degrades system performance. We apply both mixture PLDA (mix-PLDA) and trial-based calibration with condition PLDA similarity (TBC-CPLDA) to improve system robustness. Our proposed approaches resulted in 18% and 33% relative improvement in discrimination and calibration performance respectively on the SRI-FRTIV corpus.

Index Terms— Vocal effort, speaker embeddings, calibration, speaker recognition, condition PLDA

1. INTRODUCTION

Variability in the acoustic signal is a persistent challenge for speaker recognition systems operating under real-world conditions. Such variability is caused by either intrinsic or extrinsic factors. Intrinsic factors are associated with the speaker rather than the recording environment. These factors include changes in vocal effort, speaking style [1], non-speech sounds [2, 3, 4], emotions, language [5], aging, etc. across recordings of the same speaker. Extrinsic factors are associated with the differences in the recording environments between recordings. These factors include changes in background noise, microphone, room acoustics, distance from the microphone [6], transmission channel, codec [7], etc. Intrinsic factors are also known as speaker-dependent factors, whereas extrinsic factors are called speaker-independent factors [8].

During recent decades, US government evaluations and programs (such as the NIST Speaker Recognition Evaluations (SRE), the IARPA BEST program, and the DARPA RATS program) have motivated particular research directions in the speaker recognition community. Those research programs have primarily focused on the problem of extrinsic variability, including channel effects, transmission noise, and environmental noise. Intrinsic variability, in contrast, has received sparse research exposure. Yet, intrinsic variability is a key factor for unconstrained applications, such as forensic speaker recognition. This work is focused specifically on vocal effort variations, which is one class of intrinsic variability.

Vocal effort has been shown to impact the performance of speaker recognition systems [9]. In the past, a number of studies focused on different levels of vocal effort, such as whisper [10], shouts [11], and screams [4]. The impact of Lombard speech on the performance of speaker verification system was considered in [12, 13].

The main contributions of this work are as follows. First, we use a state-of-the-art DNN speaker embeddings based speaker recognition system over classical GMM-UBM or i-vector based systems. Second, rather than focusing on just one type of vocal effort level such as whisper or shouts, we develop our mitigation approaches for a range of vocal efforts from low to high. Third, we use a relatively large number of speakers with sufficient audio data per speaker to get significant results. Also, to the best of our knowledge, this study is the first to consider calibration of speaker recognition system for a range of vocal efforts.

In this study, we first assess the impact of vocal effort on discrimination and calibration performance of a DNN speaker embeddings speaker recognition system. We then apply mixture PLDA (mix-PLDA) using meta information and the recently proposed trial-based calibration with condition PLDA similarity (TBC-CPLDA) to mitigate the impact of vocal effort. We used SRI-FRTIV corpora for all the experiments.

2. CORPUS

The SRI-FRTIV (Five-way Recorded Toastmaster Intrinsic Variation)[1] corpus was collected under controlled conditions for intrinsic variability analysis at SRI International [1]. This corpus has 34 native speakers of North American English. Each speaker was recorded for eight different conditions (labeled conditions 1–8 in Table 1), which are combinations of vocal effort and speaking styles. During data collection, the unnatural combinations of speaking style and vocal effort (labeled NA) were excluded (e.g., oration at low vocal effort). Each speaker was recorded during two sessions, separated by an average of two to three weeks.

Table 1. Different conditions (1–8) in the SRI-FRTIV corpus, where NA indicates an unnatural condition.

	Normal Effort	Low Effort	High Effort
Interview	1	2	NA
Conversation	3	4	NA
Reading	5	6	7
Oration	NA	NA	8

Some unique aspects of the SRI-FRTIV corpus include: (i) the use of furtive or very low vocal effort speech as opposed to whis-

The research by authors at SRI International was funded through a development contract with Sandia National Laboratories (SNL) (Subcontract# 1758993/ DO 1872160). The views herein are those of the authors and do not necessarily represent the views of the funding agencies.

[1]https://catalog.ldc.upenn.edu/LDC2017S18

pered speech, (ii) high vocal effort being associated with projection over a distance rather than over noise (as in the Lombard effect), and (iii) the subject's position relative to the microphone being held constant across all conditions.

In this study, for the vocal-effort experiments, we use normal effort (C5), low effort (C6), and high effort (C7) conditions with a reading speaking style. For a controlled analysis of vocal effort variations, we keep all the parameters constants which are known to impact the performance of a speaker recognition system. We use read speaking style, fixed duration of speech (15 sec.) and clean speech. The 15 sec. audio files were chunked based on the output of a speech activity detection system.

For this study, we split the SRI-FRTIV data by speaker into evaluation and development or calibration sets. Each set contains recordings from 17 speakers.

3. SPEAKER RECOGNITION SYSTEM

In this section, we detail our speaker recognition system used in this work. We use a state-of-the-art DNN speaker embeddings-based speaker recognition system. The main components of our system include speech activity detection (SAD), a DNN-based embedding extractor, a probabilistic linear discriminant analysis (PLDA) classifier, and a score-calibration module.

3.1. Speech Activity Detection

In our previous work [14], we investigated the impact of speech activity detection (SAD) on the performance of speaker embeddings-based speaker recognition systems. It was shown that a low SAD threshold during training tended to benefit the embeddings extractor, while maintaining a strict threshold during evaluation was necessary. Here, our SAD is DNN-based system with two hidden layers containing 500 and 100 nodes, respectively. The SAD DNN is trained using 20-dimensional mel-frequency cepstral coefficients (MFCC) features, stacked with 31 frames. Before training the SAD DNN, the features are mean and variance normalized over a 201-frame window. The threshold for selecting the speech versus non-speech frames is 0.5 for evaluation and -1.5 for DNN training. The SAD model is trained on clean telephone and microphone data from the Mixer datasets released under the NIST SREs.

3.2. Speaker Embedding Extractor

The architecture of our speaker embeddings extractor DNN follows the Kaldi recipe [15, 16]. This feed-forward DNN is trained to discriminate between speakers. By using a statistics pooling layer, the DNN maps a variable-length utterance to a fixed-dimensional embedding. The embeddings network has five frame-level hidden layers with rectified linear unit (ReLU) activation and batch normalization. The first three layers incrementally add time context with stacking of [-2, -1, 0, 1, 2], [-2, 0, 2], and [-3, 0, 3] instances of the input feature frame. Means and standard deviations of the frame-per-audio segments are stacked using a statistic pooling layer. The final two hidden layers of 512 nodes operate at the segment level prior to the log-soft-max output layer. A ReLU activation function and batch normalization prior to a layer's output is applied to all layers except the output layer. Speaker embeddings can be extracted either from first or second segment-level hidden layer, each being 512 nodes. In this work, we use first segment-level hidden layer for extraction of the speaker embeddings [15, 17].

For the training of the embedding extractor, we followed the recipe in [14] for the raw+CNlowRM system. We used raw PRISM data, along with four types of degradation to train the embedding extractor.

3.3. Probabilistic Linear Discriminant Analysis (PLDA) Classifier

A PLDA model learns to separate within-class and across-class variability from a large, labeled training set using expectation maximization (EM) [18]. We use gender-independent PLDA for all our experiments described herein. Before training the PLDA classifier, the dimensions of the embeddings are reduced to 200 using linear discriminant analysis (LDA), followed by length normalization and mean centering [19]. Finally, these normalized speaker embeddings are used by the PLDA classifier to compute a similarity score between speaker embeddings. The full PRISM training lists (including original degradations) with additional transcoded data are used for training the PLDA model [20].

3.4. Score Calibration

The output scores of speaker recognition systems are not directly interpretable as stand-alone values. To use the speaker recognition system output scores, a calibration step is performed. The calibration step converts the system scores into meaningful output, known as log-likelihood ratios (LLRs) [21]. The LLRs have a clear probabilistic interpretation and can be either used directly in some applications, like forensic voice comparison, or converted to binary decisions by applying a score threshold for other applications, such as user authentication. We use a linear calibration transformation in which raw scores s are transformed into calibrated scores s_c, given scaling and offset parameters α and β:

$$s_c = \alpha s + \beta \tag{1}$$

where α and β are obtained by logistic regression optimization on development data.

We use the equal error rate (EER) percentage and a minimum decision cost function (minDCF) from NIST SRE 2010 evaluation with a probability of a target trial (P_{target}) equal to 0.01 and a cost of log-likelihood ratio C_{llr} to measure the performance of the system. C_{llr} measures discrimination and calibration over all possible operating points.

We use both the EER and minDCF to help draw confident conclusions when limited errors at a given operating point resulted in a wide confidence margin around one of the metrics. This may occur when too few errors or unique speakers are used in the calculation of the metric.

4. IMPACT OF VOCAL EFFORT MISMATCH

In this section, we assess the sensitivity of speaker recognition on discrimination and calibration performance to the different levels of vocal effort on SRI-FRTIV.

4.1. Discrimination Performance

For impact assessment of vocal effort, we selected three vocal effort modes (C5–C7) for the read condition from SRI- FRTIV. Audio segments were chunked into 15 seconds of speech based on decisions from speech activity detection. We enrolled speakers using normal vocal effort data from session 1, and verification was done using

different vocal efforts from session 2. We generated trials by pooling trials from all the microphones, including the cross-mic trials. Speakers models were enrolled using a single enrollment segment of 15 second, however, we enrolled each unique speaker multiple time using all available segments from session 1 to produce an extended set of trials.

Table 2. Impact of vocal effort on discrimination performance for speaker recognition on the SRI-FRTIV corpus.

Enroll-Test	Tgt/Imp	EER (%)	minDCF
Normal-Low	13.8k/227.6k	4.08	0.259
Normal-Normal	29.9k/487.7k	0.61	0.042
Normal-High	31.8k/517.5k	1.96	0.098
Normal-All	75.6k/123.2k	2.03	0.11

Results from the experiments are shown in Table 2. We observe that test speech with low vocal effort causes a significant degradation, more than twice that of normal vocal effort. Even though speaker embeddings generalize well across unseen conditions [14, 16], results in Table 2 show the pooled test condition (Normal-All) is closely tied to the performance of the Normal-High subset, rather than being similar to the average across conditions. This indicates that calibration across conditions is particularly poor for the speaker embedding system when coping with variation in vocal effort from speakers.

4.2. Calibration Performance

We assessed the impact of different vocal efforts on calibration performance for speaker recognition. For this experiment, our evaluation and calibration sets were homogeneous (i.e., just one type of vocal effort available in the evaluation and calibration sets). Benchmarking such that there is homogeneity across the calibration and evaluation sets provides a fundamental reference point to assess the impact of using a global calibration model or a model unaware of the specific conditions expected during system use.

Table 3. Impact of vocal effort in terms of C_{llr}.

Calibration Set	**Evaluation Set**		
	Low	Normal	High
Low	0.269	0.146	0.130
Normal	0.844	0.026	0.031
High	0.693	0.036	0.032

The results for vocal effort impact assessment are summarized in Table 3. A major finding is that low vocal effort poses a significant problem as opposed to normal and hard vocal efforts. For this set, also observe that the normal and high vocal effort are very close to each other in terms of calibration performance.

5. COMPENSATION OF VOCAL EFFORT MISMATCH

In this section, we describe the approaches used to mitigate the effects of vocal effort in our speaker recognition system. We first detail mixture PLDA, and then trial-based calibration with condition PLDA similarity is presented.

5.1. Mixture PLDA

We developed a mitigation approach for vocal effort that involved a meta extractor to predict the class of vocal effort (low, normal, or high) and used this information in a mixture PLDA (mix-PLDA) model, which has a PLDA model for each class, and weights each appropriately in the evaluation of a trial. These class-dependent PLDA models are trained simultaneously using an expectation-maximization (EM) algorithm. During training phase, an embedding is assigned to a particular PLDA model based on the posterior probability of vocal effort level. This approach was intially proposed for SNR-dependent PLDA in [22].

The mixture of PLDA models each correspond to a particular vocal effort level ranging from low to high. As a result, the speaker recognition system is expected to handle a range of vocal effort levels. We train our meta extractor on calibration set of SRI-FRTIV. During the evaluation phase for a speaker embedding, the marginal likelihoods from different PLDA models (specific to each vocal effort in training) are linearly combined based on the posterior probabilities of the test utterance and the enrollment utterance originating from each of the vocal effort levels. Final verification scores are the ratio of the marginal likelihoods.

Table 4. Impact of vocal effort on discrimination performance for speaker recognition on the SRI-FRTIV corpus.

	Regular-PLDA		**mix-PLDA**	
Enroll-Test	EER (%)	minDCF	EER (%)	minDCF
Normal-Low	4.08	0.259	**3.47**	**0.210**
Normal-Normal	0.61	0.042	0.61	0.042
Normal-High	1.96	0.098	**1.62**	**0.085**
Normal-All	2.03	0.11	**1.66**	**0.093**

Table 4 indicates the performance of mix-PLDA that was trained to be aware of the three vocal effort levels. It can be observed that mix-PLDA provided similar performance to Regular-PLDA when the test conditions were constrained to normal vocal effort level. It is when a variety of vocal efforts are evaluated (Normal-All), that the benefit of mix-PLDA becomes clear, resulting in more than 18% and 15% relative gain in EER and minDCF over Regular-PLDA, respectively.

5.2. Trial-Based Calibration

A novel approach for calibration, called trial-based calibration (TBC), was first proposed in [23]. This approach is similar to the way in which a forensic expert calibrates each trial individually. Trial-based calibration trains a separate calibration model for each test trial using data that is dynamically selects from a candidate training set to closely match the conditions of the trial. The model trained for each trial is thus targeted toward the characteristics of the test trial and is not contaminated by data that is markedly different from the trial conditions. A dynamic calibration model is trained with the scores of these selected trials using linear logistic regression, as is done in the case of global calibration.

A crucial factor in the TBC approach is how to determine which data, if any, is sufficiently similar to the conditions of the trial to generate an appropriate calibration model. This selection is done based on a similarity metric between the trial samples and the candidate calibration samples. In the first formulation of TBC, similarity was determined using the universal audio characterization (UAC) vector

approach [24], developed under the IARPA BEST program. This approach uses i-vectors from a speaker recognition system to train a Gaussian back-end to detect all known relevant conditions affecting SID, such as channels, reverberation times, SNRs, codec, noise type, etc. The collected scores for all the conditions form the UAC vector, which represents the meaningful condition variability of the file in a single array. When compared with UAC vectors from the candidate calibration set, the most relevant calibration data could be located.

5.2.1. Condition PLDA Similarity

Our more recent work has moved away from this approach [25], in which conditions were treated rather independently, and has instead moved toward using a common subspace that is rich in condition variability. SRI developed a new approach, condition probabilistic linear discriminative analysis (CPLDA), which more accurately compares conditions across files and improves the ability of TBC to generalize to unseen conditions. The CPLDA similarity used in this work is given by the score produced by a PLDA model trained to estimate the log-likelihood ratio of the sample's speaker embedding, given the hypothesis that they come from the same condition versus the hypothesis that they come from different conditions. The CPLDA model is trained with data from many different speakers under many different conditions, while modeling the condition variation instead of speaker variation as is done in traditional PLDA for SID [25]. For this work, our CPLDA training set consists of our PLDA list and calibration list of SRI-FRTIV with conditions labels for channel, language, vocal effort and gender. The label for the condition of an audio file is formed through the combination of these three categories.

5.2.2. Regularization of Calibration Model

In this work, we also consider a regularized version of linear logistic regression, in which a term is added to the objective function to penalize the distance from the estimated parameters to a default set of parameters. This was first proposed in [25] as well.

We use a version that we regularize toward the default parameter values. That is, we maximize the following objective function,

$$L_R(\alpha,\beta) = L(\alpha,\beta) + \lambda L_0[\frac{(\alpha-\alpha_0)^2}{\alpha_0^2} + \frac{(\beta-\beta_0)^2}{\beta_0^2}] \quad (2)$$

where L is the standard logistic regression objective function, and parameters α_0 and β_0 are learned from the pooled calibration data as would normally occur when training a global calibration model. The value of λ is chosen empirically to optimize calibration performance (0.05 was used in this work).

Table 5. Effect of regularization on vocal effort.

	Enroll-Test	Regularization	C_{llr}
TBC-CPLDA	Normal-All	No	0.455
TBC-CPLDA	Normal-All	Yes, α=0.05, β=0.05	0.130

Table 5 shows a comparison of results when using no regularization and using regularization on the trial-based calibration with condition PLDA (TBC-CPLDA) method. We observe that regularization offers a solid calibration performance improvement for vocal effort. We found that regularization parameters, alpha and beta, yielded the best results with a value equal to 0.05.

Typically, data availability is lacking for calibrating under conditions with intrinsic variability. Therefore, the TBC system can be tailored for the lack of data by setting the user-defined parameter for maximum number of target trials (MaxTgt) to a relatively small value. The regularization parameters with value equal to 0.05 allow for a MaxTgt in the range of 30 to 100 for the SRI-FRTIV dataset. For interested readers, additional details on the selection algorithm can be found in Section III.B of [25].

5.2.3. TBC-CPLDA Results

We present the results for the global calibration, and trial-based calibration with condition PLDA (TBC-CPLDA) in Table 6 and regularization. In the case of vocal effort, TBC-CPLDA is very effective.

Table 6. Mitigation of calibration of vocal effort.

Calibration	Enroll-Test	C_{llr}
Global Calibration	Normal-All	0.183
Trial-Based Calibration with CPLDA	Normal-All	0.130

We observed 29% relative improvement in C_{llr} with TBC-CPLDA compared to global calibration. The TBC-CPLDA approach outperforms the global calibration because of its ability to select the most relevant data for calibration using condition PLDA. These results indicate the degree of impact that varying vocal effort levels can have on speaker recognition calibration performance and the need to appropriately counteract such variation through robust calibration methods such as TBC-CPLDA.

6. CONCLUSIONS

This study demonstrated that variations in vocal effort level significantly degrade discrimination and calibration performance of a state-of-the-art speaker recognition system. Low vocal effort poses a significant challenge compared to high vocal effort. We used mixture PLDA to improve discrimination performance under varying levels of vocal effort. This approach leveraged an automatic estimate of vocal effort level to appropriately weight a mixture of PLDA models trained for each vocal effort condition. For robust calibration of a speaker verification system under different vocal efforts, we used trial-based calibration (TBC) with condition-PLDA. Future work will include applying the proposed approaches to a large dataset and to a wider scale of vocal effort variations, such as from whispering to shouted speech.

7. REFERENCES

[1] E. Shriberg, M. Graciarena, H. Bratt, A. Kathol, S. S. Kajarekar, H. Jameel, C. Richey, and F. Goodman, "Effects of vocal effort and speaking style on text-independent speaker verification," *INTERSPEECH-2008*, pp. 609–612, 2008.

[2] M. K. Nandwana and J.H.L. Hansen, "Analysis and identification of human scream: implications for speaker recognition," *INTERSPEECH-2014*, pp. 2253–2257, 2014.

[3] M. K. Nandwana, H. Bořil, and J.H.L. Hansen, "A new front-end for classification of non-speech sounds: a study on human whistle," *INTERSPEECH-2015*, pp. 1982–1986, 2015.

[4] J.H.L. Hansen, M. K. Nandwana, and N. Shokouhi, "Analysis of human scream and its impact on text-independent speaker

verification," *The Journal of the Acoustical Society of America*, vol. 141, no. 4, pp. 2957–2967, 2017.

[5] A. Misra and J.H.L. Hansen, "Modelling and compensation for language mismatch in speaker verification," *Speech Communication*, vol. 96, pp. 58–66, 2018.

[6] M. K. Nandwana, J. van Hout, M. McLaren, A. Stauffer, C. Richey, A. Lawson, and M. Graciarena, "Robust speaker recognition from distant speech under real reverberant environments using speaker embeddings," *INTERSPEECH-2018*, pp. 1106–1110, 2018.

[7] M. McLaren, V. Abrash, M. Graciarena, Y. Lei, and J. Pesán, "Improving robustness to compressed speech in speaker recognition," *INTERSPEECH-2013*, pp. 3698–3702, 2013.

[8] J.H.L. Hansen and T. Hasan, "Speaker recognition by machines and humans: A tutorial review," *IEEE Signal processing magazine*, vol. 32, no. 6, pp. 74–99, 2015.

[9] D. S. Brungart, K. R. Scott, and B. D. Simpson, "The influence of vocal effort on human speaker identification," *EUROSPEECH-2001*, pp. 747–750, 2001.

[10] X. Fan and J.H.L. Hansen, "Speaker identification within whispered speech audio streams," *IEEE transactions on audio, speech, and language processing*, vol. 19, no. 5, pp. 1408–1421, 2011.

[11] R. Saeidi, P. Alku, and T. Bäckström, "Feature extraction using power-law adjusted linear prediction with application to speaker recognition under severe vocal effort mismatch," *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, vol. 24, no. 1, pp. 42–53, 2016.

[12] J.H.L Hansen and V. Varadarajan, "Analysis and compensation of Lombard speech across noise type and levels with application to in-set/out-of-set speaker recognition," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 2, pp. 366–378, 2009.

[13] F. Kelly and J.H.L. Hansen, "Evaluation and calibration of Lombard effects in speaker verification," *IEEE Spoken Language Technology Workshop (SLT)*, pp. 205–209, 2016.

[14] M. McLaren, D. Castan, M. K. Nandwana, L. Ferrer, and E. Yilmaz, "How to train your speaker embeddings extractor," *Odyssey 2018*, pp. 327–334, 2018.

[15] D. Snyder, D. Garcia-Romero, D. Povey, and S. Khudanpur, "Deep neural network embeddings for text-independent speaker verification," *INTERSPEECH-2017*, pp. 999–1003, 2017.

[16] D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur, "X-vectors: Robust DNN embeddings for speaker recognition," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5329–5333, 2018.

[17] M. K. Nandwana, M. McLaren, D. Castan, J. van Hout, and A. Lawson, "Analysis of complementary information sources in the speaker embeddings framework," *INTERSPEECH 2018*, pp. 3568–3572, 2018.

[18] S. J. D. Prince and J. H. Elder, "Probabilistic linear discriminant analysis for inferences about identity," *IEEE 11th International Conference on Computer Vision (ICCV)*, pp. 1–8, 2007.

[19] D. Garcia-Romero and C. Y. Espy-Wilson, "Analysis of i-vector length normalization in speaker recognition systems," *INTERSPEECH 2011*, pp. 249–252, 2011.

[20] L. Ferrer, H. Bratt, L. Burget, H. Cernocky, O. Glembek, M. Graciarena, A. Lawson, Y. Lei, P. Matejka, O. Plchot, et al., "Promoting robustness for speaker modeling in the community: the PRISM evaluation set," *Proceedings of NIST 2011 workshop*, 2011.

[21] N. Brümmer and J. Du Preez, "Application-independent evaluation of speaker detection," *Computer Speech & Language*, vol. 20, no. 2-3, pp. 230–275, 2006.

[22] Man-Wai Mak, "SNR-dependent mixture of PLDA for noise robust speaker verification," *INTERSPEECH-2014*, pp. 1855–1859, 2014.

[23] M. McLaren, A. Lawson, L. Ferrer, N. Scheffer, and Y. Lei, "Trial-Based Calibration for speaker recognition in unseen conditions," *Odyssey 2014*, pp. 19–25, 2014.

[24] L. Ferrer, L. Burget, O. Plchot, and N. Scheffer, "A unified approach for audio characterization and its application to speaker recognition," *Odyssey 2012*, pp. 317–323, 2012.

[25] L. Ferrer, M. K. Nandwana, M. McLaren, D. Castan, and A. Lawson, "Toward fail-safe speaker recognition: Trial-Based Calibration with a reject option," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 1, pp. 140–153, 2019.