ICDT 2020

The Fifteenth International Conference on Digital Telecommunications

ISBN: 978-1-61208-768-9

February 23 - 27, 2020

Lisbon, Portugal

ICDT 2020 Editors

Pascal Lorenz, University of Haute-Alsace, France

ICDT 2020

Foreword

The Fifteenth International Conference on Digital Telecommunications (ICDT 2020), held between February 23-27, 2020 in Lisbon, Portugal, continued a series of special events related to telecommunications aspects in multimedia environments. The scope of the conference was to focus on the lower layers of systems interaction and identify the technical challenges and the most recent achievements.

High quality software is not an accident; it is constructed via a systematic plan that demands familiarity with analytical techniques, architectural design methodologies, implementation polices, and testing techniques. Software architecture plays an important role in the development of today's complex software systems. Furthermore, our ability to model and reason about the architectural properties of a system built from existing components is of great concern to modern system developers.

Performance, scalability and suitability to specific domains raise the challenging efforts for gathering special requirements, capture temporal constraints, and implement service-oriented requirements. The complexity of the systems requires an early stage adoption of advanced paradigms for adaptive and self-adaptive features. Online monitoring applications, in which continuous queries operate in near real-time over rapid and unbounded "streams" of data such as telephone call records, sensor readings, web usage logs, network packet traces, are fundamentally different from traditional data management. The difference is induced by the fact that in applications such as network monitoring, telecommunications data management, manufacturing, sensor networks, and others, data takes the form of continuous data streams rather than finite stored datasets. As a result, clients require long-running continuous queries as opposed to one-time queries. These requirements lead to reconsider data management and processing of complex and numerous continuous queries over data streams, as current database systems and data processing methods are not suitable. Event stream processing is a new paradigm of computing that supports the processing of multiple streams of event data with the goal of identifying the meaningful events within those streams.

We take here the opportunity to warmly thank all the members of the ICDT 2020 technical program committee, as well as all the reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and effort to contribute to ICDT 2020. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

We also thank the members of the ICDT 2020 organizing committee for their help in handling the logistics and for their work that made this professional meeting a success. We hope that ICDT 2020 was a successful international forum for the exchange of ideas and results between academia and industry and to promote further progress in the domain of digital telecommunications. We also hope that Lisbon, Portual provided a pleasant environment during the conference and everyone saved some time to enjoy the historic charm of the city

ICDT 2020 Chairs:

ICDT Steering Committee
Constantin Paleologu, University Politehnica of Bucharest, Romania
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Ioannis Moscholios, University of Peloponnese - Tripolis, Greece

ICDT 2020

Committee

ICDT Steering Committee

Constantin Paleologu, University Politehnica of Bucharest, Romania
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Ioannis Moscholios, University of Peloponnese - Tripolis, Greece
Sathiamoorthy Manoharan, University of Auckland, New Zealand
Bernd E. Wolfinger, University of Hamburg, Germany
Stan McClellan, Texas State University - San Marcos, USA

ICDT Industry/Research Advisory Committee

Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan
Scott Trent, IBM Research – Tokyo, Japan

ICDT 2020 Technical Program Committee

Arsalan Ahmad, National University of Sciences and Technology, Islamabad, Pakistan / Trinity College Dublin, Ireland
Ayad Al-Adhami, Plymouth University, UK / University of Technology, Iraq
Babak Barazandeh, University of Southern California, USA
Ilija Basicevic, University of Novi Sad, Serbia
Larbi Boubchir, University of Paris 8, France
Abhishek Das, Aliah University, Kolkata, India
Somaieh Davar, ConcordiaUniversity, Canada
Tan Do Duy, Ho Chi Minh City University of Technology and Education, Vietnam
Mário Ferreira, University of Aveiro, Portugal
Rita Francese, Università di Salerno, Italy
Felix J. Garcia-Clemente, University of Murcia, Spain
Benedict R. Gaster, University of West of England, UK
Carlos Guerrero, University of Balearic Islands, Spain
Onur Günlü, Technical University of Berlin, Germany
Yishan Jiao, Pearson Education, USA
Wen-Hsing Lai, National Kaohsiung University of Science and Technology, Taiwan
Jan Lansky, University of Finance and Administration, Czech Republic
Moonjin Lee, Korea Maritime and Ocean University / University of Science & Technology Korea / Research Institute of Ship and Ocean engineering, Korea
Shunbo Lei, University of Michigan-Ann Arbor, USA
Isaac Lera, University of the Balearic Islands, Spain
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Vladimir Lyashev, Huawei Technologies Co. Ltd. - Moscow Research Center,Russia
Min Ma, Google speech, USA
S. Manoharan, University of Auckland, New Zealand

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission or reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article is does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Model for Infant Acquisition of Spoken Words Using Genetic Algorithm and Fujisaki Model

Tomio Takara
Faculty of Engineering
University of the Ryukyus / Okinawa Polytechnic College
Okinawa, Japan
e-mail: takara@ie.u-ryukyu.ac.jp

Ryoichi Eto
International Systems Development Co. Ltd.
Okinawa, Japan
e-mail: ryoichi.eto@isd.co.jp

***Abstract*—We propose a new model of speech imitation and acquisition process of infants. We regard the vowel space parameters as the articulatory gesture, such as the tongue hump position and the degree of constriction of vowels. We represent the coarticulation effect using Fujisaki's generative model of speech. We model a trial and error process of the infant's speech imitation using the Genetic Algorithm (GA). In our model, we regard "command" in the Fujisaki model as the articulatory gesture and detect it from the spectral sequence using the GA. In other words, the original phonemic target is inversely estimated as the Fujisaki's command from the phonemically ambiguous speech spectrum caused by the coarticulation. Our model simulates the hypothesis that human infants acquire the normalization (inverse estimation) skill of the coarticulation through the process of imitating spoken words. We evaluated the model in listening tests using synthesized speech. We also show that the model can represent the phenomenon of "predicted sound", which is unconsciously heard as the effect of the normalization of the coarticulation, by comparing this predicted sound with the inversely estimated sound.**

Keywords- model for acquisition of spoken word; coarticulation; Fujisaki model; genetic algorithm; vowel space parameter

I. INTRODUCTION

Human infants learn articulatory behaviors, such as the tongue hump position and the degree of constriction, which are human internal function, by controlling their speech organs and imitating other people's speech. In this manner, infants acquire mapping skills between perception and articulation of speech [1].

We think that this mapping is the most important factor of human speech communication because it means that not only the other people's internal information but their intention is known. We think also that the skill of normalization of coarticulation, which is to grasp clearly ambiguous speech caused by the coarticulation, is acquired in this imitation training process.

However, there has been no simulation research on infant's acquisition process of spoken words using perception and production models combined whereas only the motor theory has been existed [4].

In this paper, we propose a new model of this speech imitation and acquisition process of infants. We show result of listening tests on the coarticulation effect, and they can be understood by our proposed model.

We regard the vowel space parameters [2][3] as the articulatory gesture [4], such as the tongue hump position and the degree of constriction of vowels. We represent the coarticulation effect [5][6] using Fujisaki's generative model of speech (Fujisaki model) [7]. We model a trial and error process of the infant's speech imitation using the Genetic Algorithm (GA) [8]. The papers [5] and [6] used the Fujisaki model, however, were not the acquisition models.

In our model, we regard "command" in the Fujisaki model as the articulatory gesture and detect it from the spectral sequence using the GA. In other words, the original phonemic target is inversely estimated as the Fujisaki's command from the phonemically ambiguous speech spectrum caused by the coarticulation.

Our model also simulates the hypothesis that human infants acquire the normalization (inverse estimation) skill of coarticulation through the process of imitating spoken words. We evaluate the model in listening tests using synthesized speech. We also show that the model can represent the phenomenon of "predicted sound", which is unconsciously heard as the effect of the normalization of the coarticulation, by comparing this predicted sound to the inversely estimated sound.

In section II, we describe our model of speech production and perception. In section III, we describe our model to imitate the speech production process. In section IV, we describe results of analysis of the coarticulation by our model and comparison of them to listening tests. Section V is the conclusion of the paper.

II. THE MODEL OF SPEECH PRODUCTION AND PERCEPTION

In this section, we describe parts of the proposed model of speech production and perception: speech analysis synthesis system, the vowel space parameter, Fujisaki model, and the genetic algorithm.

 ISBN: 978-1-61208-768-9

A. Speech analysis synthesis system

Spectral envelopes are extracted by the improved cepstral method [9] with sampling frequency of 10 kHz, frame length of 25.6ms and frame shift of 10ms. Speech is synthesized using the Log Magnitude Approximation (LMA) filter [10].

B. The vowel space parameter

Principal vectors of the principal component analysis [11] are calculated from many log amplitude spectra, which we call simply "spectra", of isolated vowels. Spectrum at each frame of a spoken word is transformed to components on the axis of principal vectors. We call this space constructed by the principal vectors as the vowel space and the components on the principal vectors as the vowel space parameters [2][3]. The vowel space can represent effectively the space where the vowel spectra vary widely because a principal axis with lager in the spectral space. Therefore, we can reconstruct a spectrum using only the components at the principal vectors with larger eigen values. The vowel space parameters were shown to have much more compressed information than that of the cepstra and can sufficiently express the phonemic characteristics of consonants [2][3].

Figure 1 shows the distribution of the vowel space parameters of isolated vowels spoken by a Japanese male speaker. The central frames of the isolated vowels were analyzed to be log magnitude spectra. Using these spectra, the principal component analysis was performed. These spectra were analyzed using the resulting principal components. This figure is very similar to the distribution of isolated vowels at two dimensional space of the first and the second Formant frequencies. This figure also corresponds very much to the place of vowels expressed in the tongue hump position and the degree of constriction. Therefore, we can use the vowel space parameters in place of the Formant frequencies or the articulatory gesture. Formant frequency cannot be detected always correctly whereas the vowel space parameter is not and can have higher order parameters than the Formant.

C. Fujisaki model

The Fujisaki model [7] is an effective model for approximating the contour of the fundamental frequency precisely for the source model of speech synthesis.

The phrase command and one accent command of the Fujisaki model were used in our proposed model to imitate information of the source. In the model of imitating information of the articulatory position, only accent commands were used and connected to be a command pattern [5]. Fujisaki model was originally applied to the fundamental frequency and then to the Formant frequency [5]. The vowel space parameter is similar to the Formant frequency. Therefore, we can apply Fujisaki model to the vowel space parameter.

D. Genetic algorithm

The (GA) [8] is a searching algorithm which simulates biological evolution. *N* individuals with chromosomes made by random numbers are generated first and they become the first generation. The fitness is decoded from the chromosomes using expression for each model. We adopted the ranking and the elite strategies in the selection mode. The cross over and mutation are carried out and they become a new generation. This algorithm was repeated to be 10th generation in this study. We adopted the Real number Coded Genetic Algorithm (RCGA) using the chromosomes with the real number.

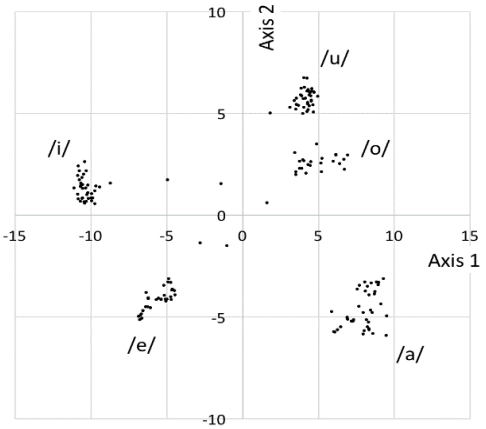

Figure 1. Analytic result of isolated vowels by the vowel space parameter

III. MODEL TO IMITATE THE SPEECH PRODUCTION PROCESS

Speech production and perception processes of this model involve four steps for each: linguistic, psychological, physiological and physical. These processes are similar to the usual speech chain, but a difference exists in the psychological process of this model. The psychological process represents one of the function of speech processing done in the human brain and is performed unconsciously. We use the vowel space parameter as the speech parameter at the psychological process. The processing of the coarticulation is performed at this process for both production and perception. We adopt the Fujisaki model as the production model for the coarticulation and the GA as the training algorithm to eliminate (or normalize) the coarticulation at the perception.

A. Acquisition of source control

1) Imitation of voiced/unvoiced decision using the GA

Figure 2 shows the coding method for voiced/unvoiced decision. The voiced or unvoiced decision is coded 1, 0, respectively. The starting time t_i is coded in the real number. First, word length is divided by number of phonemes to get equal length parts of a word. Each point with the same interval is set as an initial start point of the algorithm, then the times measured from the initial staring points are set as a starting time [×10ms] for each phoneme.

GA operation of the voiced/unvoiced decision and the fitness are as follows.

Crossover: For voiced/unvoiced decision, the code values (1 or 0) are exchanged at the crossover point. For the starting times, pairs of corresponding genes of two

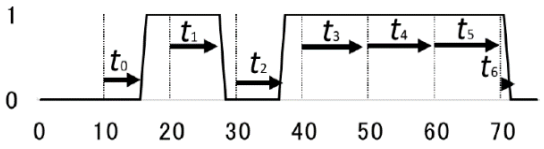

Figure 2.Coding method of voiced/unvoiced decision.

individuals are selected randomly. A new gene is generated as a random number between values of two genes. A new individual is generated by changing the gene of one of the individuals to the new gene. The other individual is similarly treated with other individual. Nine children are generated and the elite individual is added. The size of a generation was set at ten.

Mutation: For the voiced/unvoiced decision, a gene is selected with probability 0.6 for each individual and the value 1 or 0 is inversed. For the starting time, a gene is selected with a probability 0.6 for each individual and changed to be a number generated randomly between "value of the gene -20% of interval among the initial staring points" and "value of the gene +20% of interval among the initial staring points". The probability 0.6 was selected by preliminary tests.

Fitness is the percentage of the agreement of voiced/unvoiced decision for each frame.

2) Imitation of fundamental frequency using the GA

Figure 3 shows the coding method of the fundamental frequency. We adopted one phrase and one accent for the Fujisaki model. Expression of the Fujisaki model becomes as follows [7].

$$\ln F_0 = \ln F_{\min} + A_p G_p(t - T_0) + A_a\{G_a(t - T_1) - G_a(t - T_2)\} \quad (1)$$

$$G_p = \begin{cases} \alpha^2 t \exp(-\alpha t) & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (2)$$

$$G_a = \begin{cases} 1 - (1 + \beta t)\exp(-\beta t) & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (3)$$

Where F_0, $F_{\min}$, A_p, G_p, A_a, G_a, T_0, T_1, T_2, α, and β are fundamental frequency, minimal fundamental frequency, amplitude of phrase command, phrase component function, amplitude of accent command, accent component function, time point of beginning of phrase, time point of beginning of accent, time point of end of accent, time constant of phrase command function, time constant of accent command,

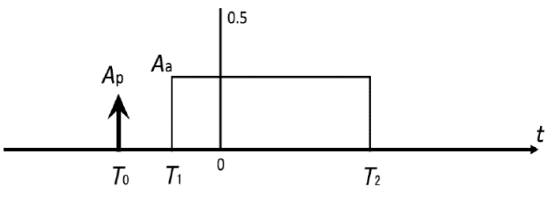

Figure 3. Coding method of fundamental frequency.

respectively. According to the preliminary test, we set $\alpha = 10$, $\beta = 25$, $A_p = 0.08$. Therefore, in order to imitate the fundamental frequency, the unknown parameters to be detected were time at phrase command T_0 and amplitude A_a, as well as beginning point T_1 and ending point T_2 of the accent command, which were coded in real numbers and were different for each word in their acquisition. The other parameters were decided using preliminary tests and set constant. The times are restricted in the conditions:

$$-20 < T_0 < 0 \quad (4)$$

$$T_0 < T_1 < T_2 < T_E \quad (5)$$

Where T_E is word length. The amplitude of accent command is

$$0 < A_a < 1. \quad (6)$$

The GA operations and the fitness of the fundamental frequency are as follows.

Crossover is the same as the above crossover of the staring time in III.*A.1 Imitation of voiced/unvoiced decision using the GA*.

Mutation is the same as the above mutation of the staring time in III.*A.1) Imitation of voiced/unvoiced decision using the GA*.

Fitness is Euclidean distance between a fundamental frequency pattern of an original word and the generated contour of the Fujisaki model, which is evaluated at the voiced interval.

3) Evaluation of the imitation model of source control

The information for source is obtained by combining information of the fundamental frequency and the voiced/unvoiced decision obtained by the GA.

In order to evaluate the obtained information of the source, we performed listening tests to check the quality of the synthesized speech made by the obtained parameters [12]. As a result of the tests, it was shown that the quality (3.2) attained was near to that of the analysis synthesis speech (3.6) at the third generation. It is very similar to human infant's linguistic performance that the model can acquire the source information in a short time with a few trials.

B. Imitation of articulatory position control [13]

The articulatory position is acquired by imitating time pattern of each dimension of the vowel space parameter extracted from speech spectrum. According to the former method applying the Fujisaki model to Formant frequencies [5], a few accent command components of the Fujisaki model are prepared to be the same number as phonemes of a word. In this expression, the ending times of phonemes are not used but the beginning times only. Therefore, the expression of the Fujisaki model for each dimension of the vowel space parameter is as follows.

$$A(t) = A_{\min} + \sum_{j=0} (A_{j+1} - A_j) G_{aj}(t - T_j) \quad (7)$$

$$G_{aj}(t) = \begin{cases} 1 - (1 + \beta t)\exp(-\beta t) & (t \geq 0) \quad (8) \\ 0 & (t < 0) \quad (9) \end{cases}$$

The time constant β was set to 20 from the preliminary test, referring to that of fundamental frequency. The unknown parameters to be detected to imitate the vowel space

parameters are A_j and T_j of the Fujisaki model, which represent characteristics of each phoneme of a word. The parameters are prepared for each phoneme and coded in a real number. The other parameters were decided using preliminary tests and set constant. The initial staring points are set similarly to the above staring time in III.*A.1) Imitation of voiced/unvoiced decision using the GA.*

The GA operations and the fitness of the vowel space parameters, which represent information of the coarticulation, are as follows.

Crossover: We used the all point crossover. Pair of corresponding genes of two individuals are selected. For each gene, a new gene is generated as a random number between values of the two genes. A new individual is generated by changing all genes of one of the individuals to the new genes. The other individual is similarly treated with other individual. Nine children are generated and the elite individual is added. The size of a generation becomes ten.

Mutation: One of the genes is selected with mutation probability 0.6 for each individual. The value 0.6 was decided by preliminary tests. The amplitude A_i is changed to be a number generated randomly between "the gene value - its 50%" and "the gene value + its 50%". The staring time T_j is changed to be the number generated randomly between "the gene value - 20% of interval among the initial staring points" and "the gene value + 20% of interval among the initial staring points".

Fitness is the Euclidian distance between the vowel space parameter obtained from the original spectrum of a word and the Fujisaki contour generated by the GA.

In order to evaluate the acquired vowel space parameters, we synthesized speech using the spectrum obtained from the imitated and acquired vowel space parameters and performed listening tests to check the quality of the synthesized speech [13]. As a result of the tests, it was shown that at the third generation, words were acquired with the quality of correct rate 90% in the listening test for vowels. It is very similar to human infant's linguistic performance that the model can acquire the articulatory position control with a few trials.

IV. ANALYSIS OF THE COARTICULATION

For spectra of a word, it was shown that the phonemic feature can be extracted as the command of the Fujisaki model from the vowel space parameter with a very ambiguous phonemic feature because of the coarticulation [13]. In this section, we discuss in more detail using listening tests and analysis by the proposed model, in which the command of the Fujisaki model represents the effect of inversely estimating (normalizing) the process of coarticulation.

Typical objects of study regarding the coarticulation are three Symmetrically Connected Consecutive Vowels [14][6][15] (3SCCV) and two Consecutive Vowels (2CV). The 3SCCV is the connected three vowels in which the beginning and the ending vowels are the same and has been used very much to analyze and model the coarticulation.

In the 2CV, there exists the phenomenon [6] that the phoneme at transitional part is not heard. For example, the transitional part of the 2CV /ai/ is acoustically [e]. Where /.../ shows that "..." are the phonemic representation. But usually humans cannot hear this transitional sound. This is also the effect of the normalization of the coarticulation.

We newly think that the following phenomena are also the effect of the normalization of the coarticulation. First, we cut the 2CV at the center of the transitional part and delete the later part. We call this sound as the Cut two Consecutive Vowels (C2CV). When we listen to the C2CV, not only the first vowel but also the second vowel is shortly heard. We think that this phenomenon shows the predicted sound which is heard because information of the second vowel exists at the former and the transitional part as the effect of the coarticulation.

A. *Three symmetrically connected consecutive vowels*

We can clearly discriminate phonemic feature of a center vowel of the 3SCCV whereas acoustic characteristics at the center do not reach the target of an isolated vowel. We think this phenomenon to be the effect of the normalization of the coarticulation in the process of human speech information processing. For example, speech /i/ in /aia/ is sometimes acoustically [e], but human hear it like [i].

We prepared the 3SCCV of /aia/, /aua/, /iui/. Figure 4 shows an example of analyzed result of the /aia/. This is the value at the third generation where the GA is performed 100 times and an individual with the best fitness at the tenth generation is selected. We set the number of phonemes to be five including the preceding and the following silence. The vertical axis shows the first dimension of the vowel space parameter. The curved solid line shows the observed vowel space parameter and the dashed line is the contour of the Fujisaki model. The straight line is the command of the Fujisaki model. The value of the command at the first /a/ is about 9, that of /i/ is about -10 and that of the last /a/ is about 8. Comparing these values to the horizontal axis of Figure 1, we can see that the value of the /i/ is grasped as the Fujisaki's command at the place acoustically [e] where the observed value of the solid line is about -7. That is to say, the model grasps the target of the phoneme while the acoustical value does not reach the target. For other dimensions of the vowel space parameter and speech data, we got almost the same results.

Figure 4 can be understood as follows. The proposed coarticulation model is already trained coefficients of the

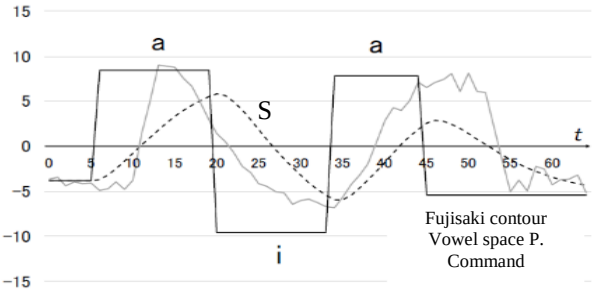

Figure 4. An example of analyzed result of three symmetrical connected consecutive vowels /aia/ by the proposed method.

expression of the Fujisaki model including the time constants. As the result, the model predicts (inversely estimates) the value of the target as the command of the Fujisaki model using the slope from the preceding phoneme to the target (S in Figure 4) under the condition of the set number of phonemes.

B. Cut two consecutive vowels

The second vowel of the C2CV doesn't exist physically but is heard auditorily. For example, when we delete the following part of /ai/ and listen, it sounds like [aj], which is usually perceived as [ai]. This phenomenon is thought to show the effect of unconscious normalization processing of the coarticulation. We study whether the proposed model of acquisition of spectral information can be a model for such a phenomenon of normalization of the coarticulation. For investigation of the C2CV, we adopted /ai/, /ia/, /au/, /ua/, /iu/, /ui/. They are the transitional phonemes among /a, i, u/ which exist in almost all languages in the world.

First, we performed listening tests to check what the inexistent vowel of the C2CVs sounds like. We used the speech data uttered three times by a Japanese male. We cut the data at 3/4 point in the transitional interval by observing the transitional contour of Formants and processed the tail by the linear fadeout with 5ms length. The beginning of the fadeout is the center of the two vowels.

All the C2CVs were listened randomly and the listeners were asked to answer in two second. The instruction to the listeners was "listen to a consecutive two vowels and answer the last vowel among /a, i, u, e, o/". The listeners were five Japanese university students with normal hearing ability. One speech specimen was used seven times and the last four results were used for the statistical analysis. Therefore, there were 60 answers for each C2CV.

Table I shows a confusion matrix of the result of the listening test. The numbers represent percentages of the perceived phonemic value, shown in the shaded block, of the inexistent vowels, shown in italic at IN for each C2CV. When we see the shaded part in the table, we find that the target vowels of C2CVs are answered at a high rate.

Figure 5 shows an example of the result of the analyzed C2CV /ai/. The vertical axis represents the first dimension of the vowel space parameter. We hypothesize to be four phonemes: silence, C2CV, silence. This is a result at the third generation which is selected as the individual with the best fitness at the tenth generation after the GA operations were repeated 100 times. We can see that the values of the command at [a] is about 9 and at [i] is about -9. Comparing to the horizontal axis of Figure 1, we can see that not only /a/ but also the value of the physically inexistent sound [i] is grasped as the Fujisaki's command. Our understanding for the mechanism of this prediction is the same as the description at the above section.

V. CONCLUSION

We proposed an infant model, which imitates and acquires spoken words. For the imitation of the source information, the fundamental frequency pattern was expressed in the Fujisaki model. For the imitation of the articulatory gesture,

TABLE I. A CONFUSION MATRIX OF THE LISTENING TEST [%].

OUT / IN	i	e	a	o	u
a_*i*	63	38	0	0	0
u_*i*	51	4	3	8	33
i_*a*	0	18	82	0	0
u_*a*	0	0	99	1	0
i_*u*	0	0	0	0	100
a_*u*	0	0	0	3	97

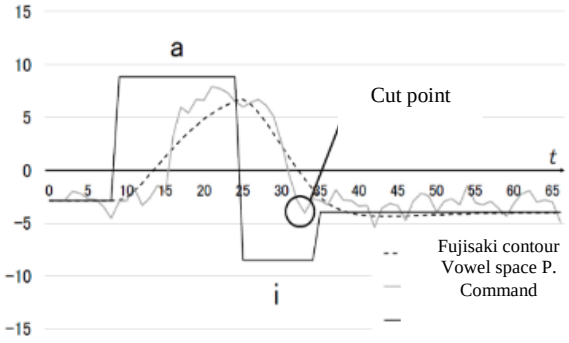

Figure 5. An example of analyzed result of the cut two consecutive vowels /ai/.

spectral information was represented in the vowel space parameter. Trial and error evident in infant's imitation process was modeled in the genetic algorithm.

We confirmed similar performance to human infants in the model of imitating spoken words, which could acquire spoken words with reasonable quality in a few trials. We showed in the analysis of the coarticulation that this model expresses suitably the normalization (inverse estimation) of the coarticulation.

This model shows that a new speech analysis method which predicts a spectrum or inversely estimates the coarticulation can be possible, especially predicts a following phoneme from the proceeding phoneme. The model also shows that a new method is possible to make a very ambiguous phonemic characteristic of spectrum in continuous speech clearer.

REFERENCES

[1] GP. K. Kuhle, B. T. Conboy, S. C. D., Padden, M. Rivera Gaxiola, and T. Nelson: "Phonetic learning as a pathway to language: new data and native language Magnet theory expanded (NLM-e)," Philos. Trans. R. Soc. B, 363, pp.979 – 1000, 2008.

[2] T. Takara, S. Yamashiro, and T. Ooishi, "Study on speech synthesis using principal components on spectral space of vowels," 2015 Annual Meeting of Acoustical Society of Japan, 2-Q-34, pp. 379-380, 2015.

[3] Tomio Takara, Akichika Higa, Syouki Kaneshiro, and Yuuya Oshiro: "Speech analysis-synthesis system using principal components of vowel spectra," The Journal of the Acoustical Society of America 140, p. 2962, 2016

[4] S. Hiroya, "Brain science of 'speaking and listening to speech': The relationship between speech production and

 ISBN: 978-1-61208-768-9

perception in brain," J. Acoust. Soc. Jpn, vol. 73, no.8, pp. 509-516, 2017.

[5] Y. Saito and H. Fujisaki, "Formulation of the process of coaticulation in terms of formant frequencies and its application to automatic speech recognition," J. Acoust. Soc. Jpn., vol. 34, pp. 177-185, 1978.

[6] M. Akagi and S. Furui, "Modeling of vowel target pre-diction mechanism in speech perception," Proc. of IECE, vol. 69A, pp. 1277-1285, 1986.

[7] H. Fujisaki and K. Hirose: "Analysis of voice fundamental frequency contours for declarative sentences of Japanese," J. Acoust. Soc. Jpn. (E) vol. 5, no. 4, pp.233 – 242, 1984.

[8] Z. Michalewicz, Genetic Algorithm + Data Structures = Evolution, Programs, 3rd ed., Springer-Verlag, Berlin, Heidelberg, pp. 97-106, 1996.

[9] S. Imai and Y. abe, "Extraction of spectral envelope using improved cepstral method," Proc. IEICE A, J62-A, 4, pp. 217-223, 1979.

[10] S. Imai, "Log magnitude approximation (LMA) filter," Proc. IEICE A, J63-A, 12, pp. 886-893, 1980.

[11] R. Schalkoff, Pattern recognition, pp. 306-307, John Wiley & Sons, Inc., 1992.

[12] T. Takara and R. Eto, "Source information for the acquisition model of spoken word using genetic algorithm and Fujisaki model," Annual meeting of ASJ autumn, 1-8-5, pp. 183 - 184, 2017.

[13] T. Takara, A. Higa, R. Eto, and Go Ishikawa, "Generative model of spectra for a word using Fujisaki model and genetic algorithm," J. Acoust. Soc. Jpn., vol. 39, no. 2, pp.147 – 149, 2018.

[14] N. Kuwahara and H. Sakai, "Normalization of coarticulation effect for a sequence of vowels in connected speech," J. Acoust. Soc. Jpn., vol. 29, no. 2, pp. 91-99, 1973.

[15] Tomio Takara and Motonori Tamaki: "A normalization of coarticulation of connected vowels using neural network," ICSLP 90 pp.1369-1372, 1990.

 ISBN: 978-1-61208-768-9

Transmission Control to Suppress Interference Between Periodic and Non-Periodic Traffic in Wireless Coexistence Scenarios

Ryota Ikeuchi and Hiroyuki Yomo
Graduate School of Engineering Science, Kansai University
3-3-35 Yamate-cho, Suita, Osaka, 564-8680 Japan
e-mail: {k097114, yomo}@kansai-u.ac.jp

***Abstract*—In this paper, we consider a wireless coexistence scenario where multi–radio platforms are employed to simultaneously support periodic and non–periodic traffic. Considering a scenario where wireless terminals generating periodic traffic over one frequency band change their operating band to the other band after detecting long–term communication failures, we consider how to suppress mutual interference between periodic and non–periodic traffic over the shared channel. In this paper, we propose a transmission control alleviating negative impact of mutual interference by exploiting interface heterogeneity, traffic periodicity, and queue management. The proposed scheme realizes high packet delivery ratio of periodic traffic by suppressing transmissions of terminals with non–periodic traffic at the timing when periodic traffic is expected to be transmitted by their hidden terminals. With computer simulations and experiments, we investigate the practicality and effectiveness of the proposed scheme.**

***Keywords*—Wireless Coexistence, Factory Automation, IEEE 802.11, IEEE 802.15.4, Internet of Things**

I. Introduction

The proliferation of diverse wireless access technologies, such as LTE, WiFi, ZigBee, Bluetooth, etc., has been accelerated during the last decade to support heterogenous traffic with different requirements. Today, we have an option to simultaneously exploit these technologies with multi–radio platforms [1][2]: for instance, small, low–price IoT devices, which are equipped with multiple interfaces operating over different frequency bands, such as 2.4/5GHz and 920MHz, are commercially available [3].

In this paper, we exploit multi–radio platforms to enhance robustness of wireless networks in a highly noisy environment. A typical use–case is factory [4], where there are many metal objects blocking communication links between transmitters and receivers [5]. Furthermore, there can be noise emitted from industrial machines, as well as interference from many radio equipment around a factory. The resulting instability of communication channels causes temporal communication failure, which can last for a long period of time. If we employ wireless devices with a single interface operating over a specific frequency band in such an unstable environment, we cannot offer reliable transmissions of data: once blocking, noise or interference is generated over an operating frequency band, each device has no way to avoid them. The lack of reliability for data transmissions in a factory can result in serious incidents that could even cause human life to be in danger. Therefore, in our work, we focus on the usage of wireless devices equipped with multiple radio interfaces operating at different frequency bands, called Flexible Terminal (FT). With FT, even if noise or interference is generated over one frequency band, its operating band can be changed to the other frequency band, which enables us to avoid communication failures due to noise and interference. More specifically, we employ radio standards operating at unlicensed frequency bands: IEEE 802.11 at 2.4 GHz and IEEE 802.15.4g at 920MHz since these standards are widely employed in many industrial fields [6].

Besides the heterogeneity of radio interface, the heterogeneity of communication traffic has become a common trend in current wireless networks. In addition to non–periodic (bursty) traffic generated by classical applications, such as Internet access and video/image transfer, more deterministic and periodic traffic has become a dominant pattern especially in a scenario with sensor devices deployed for monitoring purpose [7][8]. In general, small amount of data is generated by sensor devices, for which 920MHz radio supporting low data rate with large coverage is a favorable option. On the other hand, 2.4GHz commonly used by WiFi offers higher data rate with smaller coverage than 920MHz, which makes it suited for supporting Internet access and transfer of large–size image/video files. In this work, we employ FTs to simultaneously support periodic and non–periodic traffic. In a normal operation mode without any noise or interference, FTs with non–periodic traffic employ an interface operating at 2.4 GHz while FTs with periodic traffic use an interface operating at 920MHz. Then, we consider a scenario where noise or interference is generated by surrounding devices/machines over 920MHz, and each FT with periodic traffic changes its operating interface to that at 2.4GHz. In this case, there is mutual interference between FTs with periodic traffic and FTs with non–periodic traffic. In this work, we propose a transmission control, which suppresses mutual interference by exploiting interface heterogeneity, traffic periodicity, and queue management. In the proposed scheme, FTs with non–periodic traffic detect possible hidden FTs with periodic traffic by using difference of propagation characteristics of different frequency bands. Then, FTs with non–periodic traffic predict the transmission timing of FTs with periodic traffic, and suppress their packet transmissions at the predicted timing with adaptive queue management. With computer simulations and experiments, we investigate the practicality and possible gain of the proposed scheme.

The rest of the paper is organized as follows. After describing the system model and problem definition in Section II, we

 ISBN: 978-1-61208-768-9

present our proposed transmission control in Section III. After showing and discussing some numerical results in Section IV, Section V concludes the paper with several future work.

II. System Model and Problem Definition

In this section, we first describe the system model considered in this paper, followed by the problem formulation.

A. System Model

In this work, we employ FTs with interfaces operating at 2.4GHz and 920MHz. In general, 920MHz signals have larger propagation distance than 2.4 GHz while the former achieves lower data rate than the latter. We consider a factory–like indoor area where FTs and a single Flexible Gateway (FG), which is in charge of aggregating data generated by FTs, are deployed as shown in Figure 1. The FG is also equipped with 2.4GHz and 920MHz interfaces to receive data from FTs. Some FTs are supposed to generate non–periodic, bursty, and heavy–load traffic, which are called NP–FTs. Since this type of traffic is in general supported by higher PHY rate at 2.4GHz that has limited communication range, we assume that NP–FTs are deployed near FG. On the other hand, FTs except for NP–FTs are assumed to generate periodic, light–load traffic, which are called P–FTs. A typical example of P–FT is a sensor device generating monitoring data of industrial machines and/or a given environment, which are deployed at various places within an area. This requires P–FTs to employ an interface and/or parameters realizing a larger communication range, for which 920MHz is more favorable option. We assume that the information on period of P–FT's traffic is known and shared by all FTs/FG. This is a reasonable assumption since these terminals and gateway are considered to be deployed by a single administrator of a factory. Furthermore, the timing of packet–generations of P–FTs are controlled to be equally separated over time so that they are not overlapped. This enables us to avoid contention among P–FTs. In a normal operation mode, NP–FTs employ 2.4GHz interface while P–FTs utilize 920MHz interface. Here, 2.4 GHz interface is supposed to follow IEEE 802.11 PHY/MAC protocol while 920MHz interface is in accordance with IEEE 802.15.4g/e PHY/MAC protocol. Note that both of these standards employ CSMA/CA protocol. The FG receives data from both NP–FTs and P–FTs by using its two interfaces. It is assumed that the carrier–sense range of 2.4GHz interface is smaller than that of 920MHz as shown in Figure 1: for example, carrier–sense range of NP–FT1 in Figure 1 over 920MHz is sufficiently large to detect signals transmitted by all terminals while it can only sense signals transmitted by a part of terminals over 2.4GHz.

B. Problem Definition

In this work, we consider a scenario where severe noise/interference is caused over 920MHz, which can be emitted from industrial machines and/or radio devices deployed inside/outside a factory area, and 920MHz interface suffers from continuous communication failures for a long period of time. As mentioned in Section I, FTs are able to switch

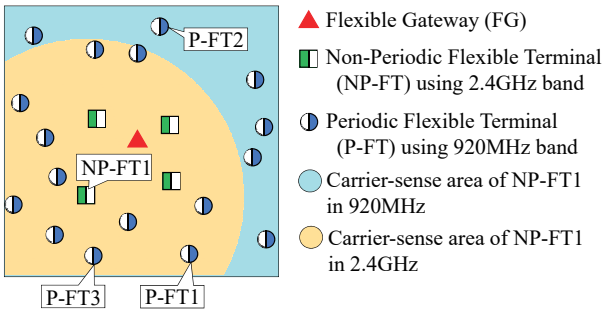

Figure 1. The considered, factory–like system model.

their operating interface. Therefore, P–FTs, which operate with 920MHz interface in a normal operation mode, can switch their operating interface to 2.4GHz, e.g., after detecting continuous packet errors or after receiving some instruction if there is a central entity to monitor the radio environment. Here, each P–FT is assumed to employ low PHY rate (e.g., 1Mbps) at 2.4GHz, which enables each P–FT to achieve sufficiently large communication range to transmit data to FG. However, when P–FTs and NP–FTs share the same 2.4GHz frequency band, another problem can occur, which is a hidden terminal problem. For example, as shown in Figure 1, NP–FT1 and P-FT2 cannot sense their signals with each other at 2.4GHz. Therefore, CSMA/CA mechanisms do not work properly among these nodes after P–FT2 changes its operating band to 2.4GHz, which can cause packet losses at FG, thereby degrading packet delivery ratio and throughput.

A well–known solution to hidden terminal problem is RTS/CTS handshake defined in IEEE 802.11. However, it has been reported that the efficiency of RTS/CTS handshake is low when short packets, such as small amount of data generated by P–FTs in our scenario, are involved in data transmissions [9]. Furthermore, RTS/CTS mechanism does not fundamentally solve problems on collisions among hidden terminals: RTS frames transmitted by hidden terminals can collide with high probability. Another requirement specific to industrial applications is more strict and deterministic protection for sensing data in comparison to Internet access/file transfer [10], which is difficult to achieve with RTS/CTS handshake even with QoS differentiation defined in IEEE 802.11e [11]. Therefore, in this work, we propose a mechanism to deterministically suppress transmissions of NP–FTs to avoid interference with hidden P–FTs without resorting to RTS/CTS mechanisms by exploiting interface heterogeneity, traffic periodicity, and adaptive queue management.

III. Proposed Transmission Control

The proposed scheme controls packet transmissions of NP–FTs in order to suppress interference with their hidden P–FTs.

A. Mechanism to Detect Hidden Terminals

The NP–FTs first need to identify possible hidden terminals in order to suppress their mutual interference. This is achieved by exploiting the heterogeneity of interface. Each NP–FT observes traffic over 920MHz and 2.4GHz while they are not

 ISBN: 978-1-61208-768-9

transmitting their own data. In the normal operation mode, P–FTs transmit data at 920MHz. In this case, each NP–FT finds packets of all P–FTs over 920MHz since they can easily reach each NP–FT thanks to a large communication range of 920MHz. For example, NP–FT1 shown in Figure 1 observes periodic receptions of all P–FTs at 920MHz interface in a normal operation mode. After P–FTs detect noise/interference at 920MHz, they switch their interface to 2.4GHz, where NP–FT1 receives packets only from P–FTs located within its communication range over 2.4GHz. Thus, in the example of Figure 1, NP–FT1 cannot receive packets transmitted by P–FT2 since P–FT2 is out of carrier–sense/communication range of NP–FT1. Then, NP–FT1 finds that it has a hidden terminal of P–FT2 over 2.4GHz. At this timing, NP–FTs can also find that P–FTs have changed their operating band to 2.4GHz. Thus, by comparing packet receptions at 920MHz and 2.4GHz, each NP–FT can identify its hidden terminals over 2.4GHz, whose packets can cause collisions against itself.

B. Basic Idea of Proposed Transmission Control

While receiving packets from P–FTs in the normal operation mode, each NP–FT records the reception timing of each P–FT. Based on this information and pre–knowledge of the period of packet transmissions of each P–FT, each NP–FT predicts the timing of periodic packet transmissions. Then, each NP–FT suppresses its packet transmissions when the transmissions of its hidden P–FTs are expected. This is achieved by our proposed Transmission Control (TC), which executes queue management to control timing to pass upper–layer packets to MAC layer module.

The basic idea of the proposed TC is shown in Figure 2. Here, the blue solid arrow shows the predicted transmission timing of a hidden P–FT. With the proposed TC, a duration called Suspending Duration (SD), which consists of Pre–SD (before the predicted timing) and Post–SD (after the predicted timing) is prepared. A NP–FT attempts to suspend its packet transmission over SD, i.e., even if packets are generated at upper layer, it stores them into its upper–layer queue, and does not pass them to MAC layer module. In Figure 2, the dashed green arrow represents the timing when packets are generated at upper layer of NP–FT. Once SD is over, NP–FT passes the stored packets to MAC layer module, which are then transmitted by MAC layer module over the air. Note that packets generated at non–SD duration can be immediately passed to MAC layer module and transmitted as in the packet P4 in Figure 2. The flowchart of these operations of the proposed transmission control is shown in Figure 3. The duration of Pre–SD and Post–SD are decided considering trade–off between achievable Packet Delivery Ratio (PDR) of P–FTs and throughput of NP–FTs as discussed in Section IV-B in more detail.

With the above operation of TC, we can prevent NP–FTs from passing their packets to MAC layer module at the timing when hidden terminals are expected to transmit their packets, thereby suppressing interference. However, the queue management at upper layer has difficulty to precisely control the timing when signals are actually transmitted at

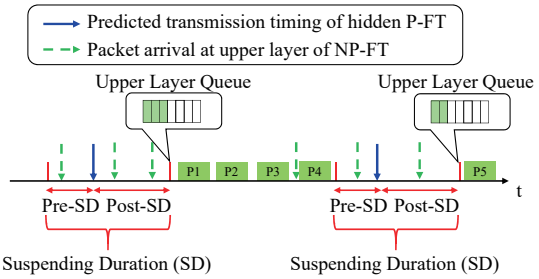

Figure 2. Basic idea of the proposed transmission control.

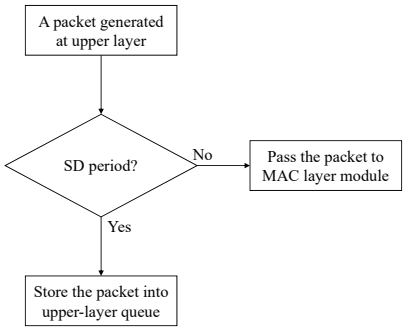

Figure 3. A flowchart of basic operations of the proposed transmission control.

PHY/MAC level. This problem is explained in an example shown in Figure 4. Here, a NP–FT suspends passing packets to MAC layer module during the first SD, and three packets are stored in the upper–layer queue. These packets are passed to MAC layer module after the first SD is over, which are then stored in lower–layer PHY/MAC queue. The transmissions of packets in PHY/MAC queue are managed by PHY/MAC module, which are in general hard to control since it requires the modification of firmware installed into WiFi module/chip. The packets are transmitted if they win contentions against the other terminals. In the example of Figure 4, it is supposed that NP–FT succeeds in transmitting a packet P1 by winning the contention. However, it fails to transmit packets P2 and P3 due to the lost contentions with BackGround (BG) traffic. Then, these 2 packets remain in PHY/MAC queue in the beginning of the next SD. As mentioned above, it is impossible to control the transmissions of these lower–layer packets, therefore, they can be transmitted even during SD, which can cause a collision with packets transmitted by hidden P–FTs.

A possible solution to the above–mentioned problem is to control the number of packets to be passed to MAC layer module based on the congestion level over the channel, i.e., each NP–FT controls the number of packets passed to MAC layer module in the end of SD in such a way that these packets can be transmitted in the following non–SD period at the PHY/MAC level. This requires each NP–FT to continuously monitor the congestion level over the operating channel. Note that background traffic at 2.4GHz are not necessarily generated by WiFi terminals, whose packets can be decoded by NP–FT, but generated by the other radio equipment, e.g., Bluetooth or Microwave oven. In this case, each NP–FT needs to monitor the congestion level without decoding each background signal.

 ISBN: 978-1-61208-768-9

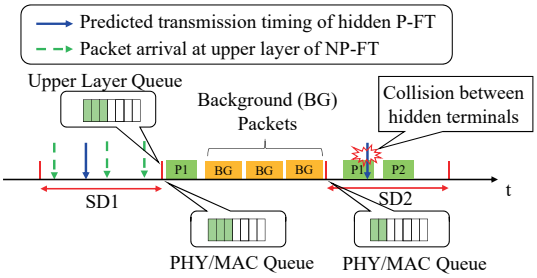

Figure 4. An example of problem on controlling packet transmissions with SD.

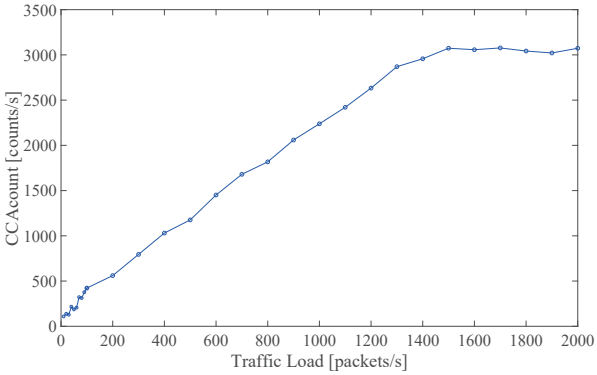

Figure 5. Experimental results on CCAcount against traffic load.

Therefore, in the following subsection, we first investigate whether it is practically possible for a WiFi terminal to conduct real–time monitoring of busy rate (i.e., fraction of time during which the channel is occupied by radio signals) of a channel.

C. Feasibility to Monitor Congestion Level

We found a parameter called CCAcount in a device driver of an off–the–shelf WiFi module (Buffalo WL–U3–866DS [12]). The parameter seems to be related to busy rate of a channel, however, there was no evidence that this parameter represents our desired information on busy rate. Therefore, we conducted experiments to check the relationship between CCAcount and busy rate of a channel. In the experiments, we prepared 3 laptop PCs with USB dongles of WL–US–866DS. A laptop PC (Tx PC) was configured to be a transmitter of packets, which are directed to Rx laptop PC. A laptop PC to observe CCAcount was located at a sufficiently close position to Tx PC. The busy rate was varied by changing the number of packets transmitted per a unit time, for which the output of CCAcount was monitored at the observing PC. The PHY rate, packet size, and ACK size of packet transmissions were respectively set to be 54Mbps, 1496Bytes, and 46Bytes. The busy rate for each traffic load can be calculated based on these parameters. The measurements were conducted inside a shielded room.

Figure 5 shows the output of CCAcount against traffic load (packets/s). From this figure, we can see that CCAcount increases as traffic load increases, which saturates over the range of high traffic load. There is a maximum traffic load that can be generated by a single WiFi terminal, which depends on back–off parameters and Inter–Frame Space (IFS) of IEEE 802.11, where the saturation is observed. From this figure, we can confirm that there is a direct relationship between CCAcount and traffic load, i.e., busy rate of the channel, which enables us to employ CCAcount as a measure of busy rate of the channel.

D. Proposed Adaptive Transmission Control

In this work, we design an Adaptive TC (ATC), which controls the number of packets to be passed to MAC layer module based on the observed CCAcount. In ATC, each NP–FT observes CCAcount during each non–SD period. The output of CCAcount is converted to the traffic load by using a linear equation approximating the relationship between CCAcount and traffic load over the load range of [0:1500] packets/s in Figure 5, which is used to calculate the busy rate. Based on the derived busy rate, the maximum number of packets permitted to be passed to MAC layer module at the next non–SD period, N_{max}, is decided. N_{max} is calculated as follows:

$$N_{max} = \frac{(1 - B_{ave})T_{NSD}}{T_D \cdot \alpha}. \quad (1)$$

Here, T_{NSD} is the duration of next non–SD period, T_D is the duration required to transmit a single data frame including SIFS and ACK duration, and α is a parameter to vary effective number of N_{max}, and B_{ave} is average busy rate calculated as

$$B_{ave} = \frac{\sum_{i=1}^{W} B_i}{W}, \quad (2)$$

where B_i is busy rate calculated for the i-th last non–SD period, and W is the window size (number of non–SDs) used for calculating average busy rate. N_{max} calculated with (1) represents the estimated (effective) number of packets that can be transmitted by a single NP–FT during free period in the following non–SD period. Note that α is introduced in order to take the impact of back–off duration and number of contending FTs into account. With smaller (larger) α, the estimation of N_{max} becomes more optimistic (pessimistic). The range of α considered in this paper is set to [0.4, 6.0].

The proposed ATC is executed in the end of every SD period. For instance, in the end of SD1 in Figure 4, N_{max} is calculated by using busy rate over the last W non–SD periods. Then, if the number of packets stored in the upper–layer queue is equal to or more than N_{max}, only N_{max} packets out of stored packets are passed to MAC layer module, and no more packets are passed to MAC layer module during the following non–SD period. Otherwise if the number of packets stored in the upper–layer queue is less than N_{max}, all stored packets are passed to MAC layer module. Then, newly arriving packets in the following non–SD period can be passed to MAC layer module as long as the total number of packets passed to MAC layer module does not exceed N_{max}. With these operations, we can reduce the probability that packets remain in PHY/MAC queue in the end of each non–SD period.

 ISBN: 978-1-61208-768-9

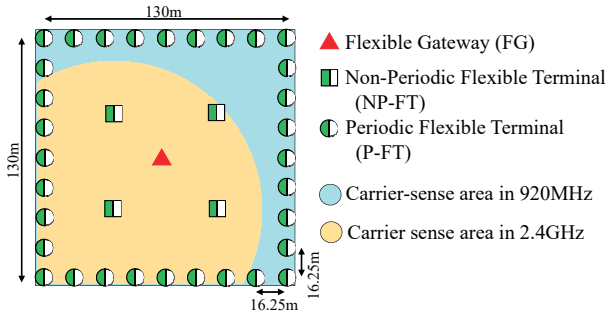

Figure 6. Simulation Model.

TABLE I: Simulation Parameters

	NP–FT	P–FT
PHY rate	54Mbps	1Mbps
Communication range	75m	100m
Carrier-Sense Range	100m	100m
Packet generation	Poisson (mean λ)	period = 1s
Data size	2000Bytes	200Bytes
ACK size	30Bytes	
DIFS	28μs	
SIFS	10μs	
Slot time	20μs	
Max. Num. of Retransmissions	3	
Min. Contention Window	31	
Simulation Duration	20s	

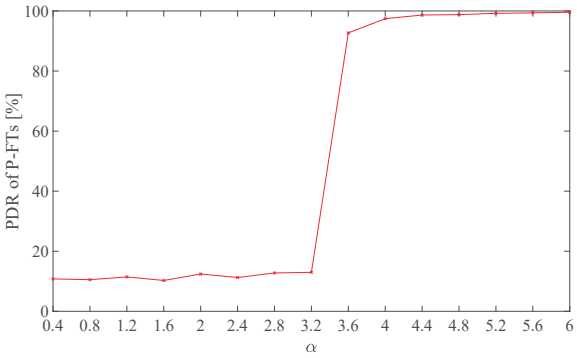

Figure 7. PDR of NP-FTs against α for ATC.

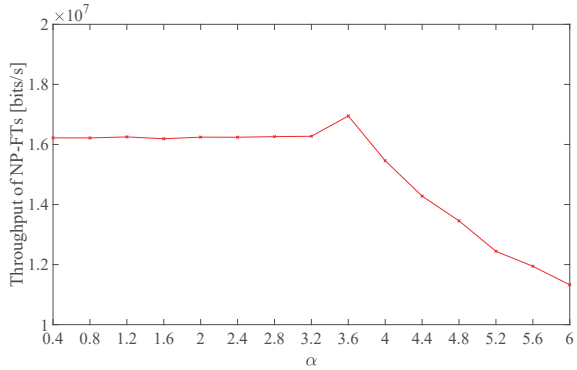

Figure 8. Throughput of NP-FTs against α for ATC.

IV. Simulation Model and Results

In this section, we provide numerical results obtained by our computer simulations, and discuss the benefit brought by the proposed transmission control in detail.

A. Simulation Model

The simulation model is shown in Figure 6. The layout given in Figure 6 is selected since it can increase the number of hidden terminals, which allows us to consider a worst–case scenario. In the simulations, communication performance after P–FTs change their operating frequency band to 2.4GHz is evaluated. The main parameters used in simulations are shown in Table I. Most of the parameters are taken from the IEEE 802.11g standard [13]. The P–FTs generate packets with period of 1s, and their generation timing are scheduled so that they do not overlap with each other. In the evaluation, since there are 32 P–FTs, a period of 1s is divided into 32 sections, and the beginning of each section is randomly assigned to each P–FT as its generation timing. Each NP–FT applies the proposed TC/ATC to its hidden P–FTs. We use the application–level PDR of P–FT and throughput of NP–FT as performance measures. A packet is decided to be lost and discarded once the number of retransmissions reaches the maximum value. For simplicity, packet errors are assumed to occur only due to collisions. The throughput is defined as the amount of data successfully delivered by NP–FTs to FG. The simulation is conducted by a custom–made simulator developed with Matlab software.

B. Simulation Results

Below, we show simulation results averaged over 5 simulation trials. Figure 7 shows PDR of P–FTs against the parameter of α in (1) when the proposed ATC is employed with Pre–SD = 2ms, Post–SD = 6ms, W = 10, and λ = 400 [packets/s]. From Figure 7, we can see that PDR of P–FT is degraded with smaller α. With smaller α, each NP–FT passes a larger number of packets to MAC layer module in the end of SD as calculated by (1), which exceeds the number of packets that can be transmitted at PHY/MAC layer during the next non–SD period. In this case, packets remained in PHY/MAC queue can be transmitted simultaneously with hidden P–FTs, which causes collisions with high probability. This problem is alleviated by increasing the value of α where the number of packets passed to MAC layer module is reduced. Therefore, PDR of P–FT is improved with larger value of α. However, larger values of α force each NP–FT to keep more packets in its upper–layer queue, and degrade its throughput performance. This is confirmed in Figure 8, where throughput of NP–FTs against α is shown. The throughput of NP–FTs is largely degraded with too large α, i.e., the range of α exceeding 3.6. From these results, we can see that there is an appropriate value of α to be employed to achieve both high PDR of P–FTs and high throughput of NP–FTs. In the following evaluations, we employ $\alpha = 3.6$ based on the above results.

Next, we investigate the impact of SD length on the achievable performance of the proposed ATC. Figures 9 and 10 respectively show PDR of P–FTs and throughput of NP–

 ISBN: 978-1-61208-768-9

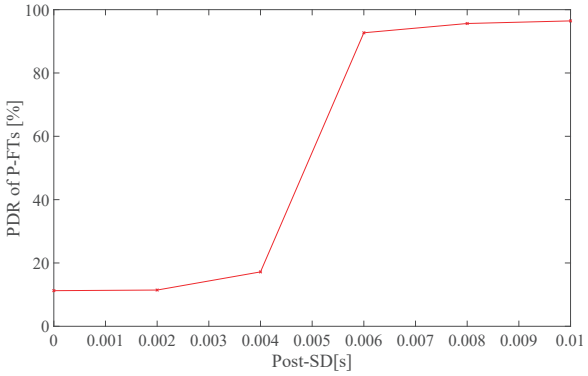

Figure 9. PDR of P-FTs against Post-SD for ATC.

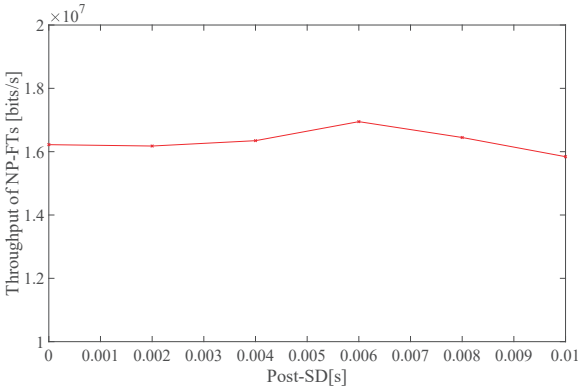

Figure 10. Throughput of NP-FTs against Post-SD for ATC.

FTs against the length of Post–SD, where Pre–SD is fixed to be 2ms, W = 10, α = 3.6, and λ = 400 [packets/s]. First, from Figure 9, we can see that a sufficiently large value of Post–SD is required to achieve high PDR of P–FTs. Each packet generated at P–FT is transmitted with CSMA/CA protocol, where its actual transmission timing at PHY/MAC level can be delayed due to contentions with the other NP–FTs and P–FTs within its carrier–sense range. Therefore, if NP–FT employs too small Post–SD, it can transmit packets with hidden P–FTs whose transmissions are delayed due to CSMA/CA operations. The increase of Post–SD also offers the improvement on throughput as shown in Figure 10 thanks to higher probability to avoid mutual collisions, however, too large Post–SD leads to the reduction of throughput of NP–FTs since it can reduce the duration for NP–FTs to be able to transmit their packets. From these figures, we can see that Post–SD of 6ms is an appropriate choice in our considered settings.

Finally, we respectively show PDR of P–FTs and throughput of NP–FTs against traffic load of NP–FTs for different schemes in Figures 11 and 12. Here, we set Pre–SD = 2ms, Post–SD = 6ms, W = 10, and α = 3.6. In these figures, we also show upper–bounds, which are obtained if we can ideally stop/start the transmission of packets at PHY/MAC level according to the schedule of SD and non–SD. Note that the performance of these upper–bounds can be obtained only if we can modify PHY/MAC module so that we can arbitrarily

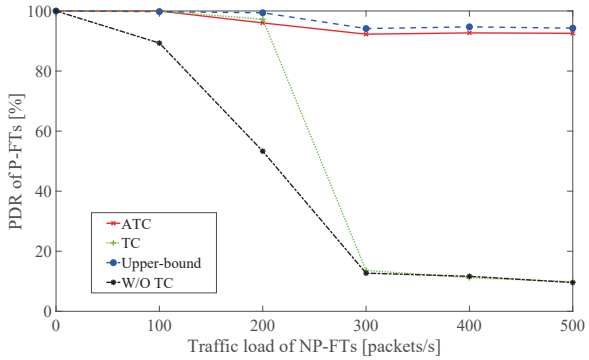

Figure 11. PDR of P-FTs against traffic load of NP-FTs.

control transmission timing at PHY/MAC level. On the other hand, our proposed ATC only requires the modification of a device driver of WiFi module. The results of W/O TC in these figures represent achievable performance of an existing scheme, which follows conventional IEEE 802.11 MAC protocol without employing our proposed TC. From Figure 11, we can first see that PDR of P–FTs is largely degraded if we do not employ TC. This is due to packet collisions between NP–FTs and their hidden P–FTs. By introducing TC with SD, PDR of P–FTs can be improved, however, we can obtain gain only over the range of small traffic load of NP–FTs. As the traffic load of NP–FTs increases, more packets are stored in the upper–layer queue in the end of each SD, which can exceed the number of packets that can be handled at PHY/MAC level during the following non–SD period. Therefore, more collisions occur for larger traffic of NP–FTs, which degrades PDR of P–FTs. On the other hand, it can be seen that the proposed ATC achieves high PDR of P–FTs even for larger traffic load of NP–FTs thanks to the adjustment of number of packets passed to PHY/MAC queue, which is adapted to the observed traffic load. We can see that the proposed ATC achieves PDR close to the upper–bound. Next, from Figure 12, we can see that the proposed ATC does not degrade throughput of NP–FTs even with the introduction of SD. The avoidance of collisions eventually leads to throughput improvement in comparison to the other schemes. With the proposed ATC, packets are stored in the upper–layer queue according to the estimated traffic load. If the actual traffic load is smaller than the estimated value, all packets passed to PHY/MAC queue can be transmitted at early timing within a non–SD period, after which no packet is transmitted since there is no packet in PHY/MAC queue. This problem does not occur with upper–bound, therefore, throughput of the proposed ATC does not reach close to the upper–bound. However, from these results, we can confirm that the proposed ATC can significantly improve PDR of P–FTs while achieving slightly better throughput of NP–FTs in comparison to the other schemes.

V. Conclusions

In this paper, focusing on a wireless coexistence scenario where multi–radio platforms are employed to support heterogenous traffic, we proposed an Adaptive Transmission

 ISBN: 978-1-61208-768-9

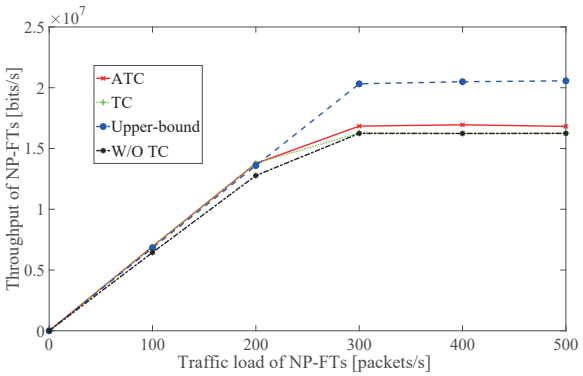

Figure 12. Throughput of NP-FTs against traffic load of NP-FTs.

Control (ATC), which suppresses mutual interference between hidden terminals generating periodic and non–periodic traffic. The proposed ATC exploits interface heterogeneity, traffic periodicity, and queue management adapting to the observed congestion level. We first confirmed with experiments the practicality for WiFi device to monitor congestion level in a real–time manner. Then, we evaluated the gain of the proposed ATC in terms of packet delivery ratio and throughput by computer simulations. Our numerical results showed that the proposed ATC significantly improves PDR of periodic traffic while slightly improving throughput of non–periodic traffic in comparison to reference schemes.

Our future work includes experimental evaluations of the proposed ATC with actual multi–radio platforms. More extensive verifications of simulation results, e.g., with a larger number of simulation trials and comparison with theoretical results, are also our future work. Furthermore, in this paper, it is assumed that the transmission timing of P–FTs can be ideally estimated by NP–FTs. However, in practice, this estimation can be incomplete, which can shift SD from the desired duration. This causes degradation of PDR of P–FTs and throughput of NP–FTs, therefore, we need to evaluate the impact of estimation error on the achievable performance of the proposed transmission control.

ACKNOWLEDGEMENT

This work is supported by the Ministry of Internal Affairs and Communications for the project entitled “R&D on Technologies to Densely and Efficiently Utilize Radio Resources of Unlicensed Bands in Dedicated Areas.”

REFERENCES

[1] J. Gummeson, D. Ganesan, M. D. Corner, and P. Shenoy, “An Adaptive Link Layer for Heterogenous Multi-Radio Mobile Sensor Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1094–1104, Sept. 2010.

[2] D. K. Tosh and S. Sengupta, “Heterogenous Access Network(s) Selection in Multi-Interface Radio Devices,” *IEEE International Workshop on Managing Ubiquitous COmmunications and Services*, pp. 117–122, 2015.

[3] “Smart Energy Gateway CUBE, NextDrive Co.” accessed: 2020-01-07. [Online]. Available: https://www.nextdrive.io/en/productNew/Cube

[4] S. Dietrich, G. May, O. Wetter, H. Heeren, and G. Fohler, “Performance indicators and use case analysis for wireless networks in factory automation,” in *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2017, pp. 1–8.

[5] M. Dungen *et al.*, “Channel measurement campaigns for wireless industrial automation,” *Automatisierungstechnik*, vol. 67, no. 1, pp. 7–28, 2019.

[6] A. Varghese and D. Tandur, “Wireless requirements and challenges in Industry 4.0,” *Proc. of 2014 International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 634–638, 2014.

[7] M. Sansoni *et al.*, “Comparison of M2M Traffic Models Against Real World Data Sets,” *Proc. of IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 1–6, 2018.

[8] X. Cao, J. Chen, Y. Cheng, X. X. Shen, and Y. Sun, “An Analytical MAC Model for IEEE 802.15.4 Enabled Wireless Networks With Periodic Traffic,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 10, pp. 5261–5273, Oct. 2015.

[9] L. Zhang *et al.*, “Signal Strength Assistant Grouping for Lower Hidden Node Collision Probability in 802.11ah,” *Proc. of 9th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–6, Oct. 2017.

[10] A. A. K. S., K. Ovsthus, and L. M. Kristensen, “An Industrial Perspective on Wireless Sensor Networks: A Survey of Requirements, Protocols, and Challenges,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1391–1412, Third Quarter 2014.

[11] “Ieee standard for information technology–local and metropolitan area networks–specific requirements–part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications - amendment 8: Medium access control (mac) quality of service enhancements,” *IEEE Std 802.11e-2005 (Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003)*, pp. 1–212, 2005.

[12] “BUFFALO, WI-U3-866DS,” accessed: 2020-01-07. [Online]. Available: https://www.buffalo.jp/product/detail/wi-u3-866ds.html

[13] D. Vassis, G. Kormentzas, A. Rouskas, and I. Maglogiannis, “The ieee 802.11g standard for high data rate wlans,” *IEEE Network*, vol. 19, no. 3, pp. 21–26, May 2005.

 ISBN: 978-1-61208-768-9