

Ministry of Higher Education & Scientific Research

University of Mosul

College of Engineering



Enhancement of Voice Transmission over Wireless Sensor Network (VoWSN) System

**Ph.D. Thesis
Submitted by**

Ina'am Fathi Khudher

Computer Engineering / Computer Networks

Supervised by

**Prof. Dr.
Qutaiba I. Ali**

**Prof. Dr.
Jassim M. Abdul-Jabbar**

2020 A.D.

1441 A.H.

Abstract

Depending on voice transmission mechanism, this thesis categorized VoWSN applications into two categories: VoWSN based on voice streaming system and VoWSN based on Automatic Speech Recognition (ASR) system. However, supporting Real-Time voice transferring using low cost, resource – limit sensor networks, is never a trivial task as it represents a demanding process in term of memory storage, CPU utilization and bandwidth utilization. So, in this thesis we aimed to deploy the two categories, analyze and compare between their performance to select the mechanism that gives the satisfied performance then suggest the suitable solutions to enhance this performance.

In order to deploy VoWSN based on streaming system, a framework to be used for processing, compressing and streaming voice was proposed and implemented. Three different audio encoding techniques were used for compressing voice data. The framework performance was evaluated using performance evaluation parameters: One-Way Delay (OWD), jitter, B.W, memory usage and CPU usage. Furthermore, to investigate the behavior of the system when connecting to the Internet, the proposed framework was connected to the Internet through Wi-Fi network to deploy Voice over Internet of Things (VoIoTs). The results showed that the OWD and jitter can keep Real-Time constrains (OWD <150 ms , Jitter <5ms) for the VoWSN and VoIoTs with different sampling frequencies.

The process of ASR with its usual requirements faces specific challenges in hardware and algorithms design when deployed using

limited resources WSN devices. Therefore, the second step was aimed to implement a high performance continuous Real-Time ASR system on WSN through using a communication procedure that based on optimizing dictionary and language models which are associated with the selected ASR system to limited required sentences. The system relied on a proposed protocol that transforms the system category from network dependent ASR system with a full dictionary and language models (i.e. large vocabulary) to fully embedded ASR system with customized dictionary and language models (i.e. small vocabulary). The results of evaluating the proposed system showed an enhancement in the Real-Time performance metrics using optimized dictionary and language models where Word Error Rate (WER) reduced about 67.11%, recognition speed enhanced about 3.96% and memory usage reduced about 8.2%. While the final step was aimed to propose a strategy for evaluating the performance of the suggested ASR system compared to the voice streaming system to explain the gains (in term of a network capacity) obtained by deploying the ASR system in the same circumstances. This step was performed by modeling the suggested platforms using OPNET simulation package. The results obtained showed a noticeable increment in the system capacity when using ASR system compared to streaming system.

الملخص

بناءً على دراسة مستفيضة للأعمال السابقة ، صنفت هذه الأطروحة تطبيقات أل VoWSN إلى فئتين اعتماداً على آلية نقل الصوت: النظام الأول هو نظام ال VoWSN المعتمد على نظام الدفع الصوتي (Voice Streaming) والنظام الثاني هو نظام ال VoWSN الذي المعتمد على نظام التعرف التلقائي على الكلام ((Automatic (ASR) Speech Recognition System)). بصورة عامة، فإن تنفيذ النقل الصوتي في الزمن الحقيقي باستخدام شبكات استشعار منخفضة التكلفة ومحدودة الموارد، لا يعتبر مهمة سهلة لأنه يمثل عملية متطلبة من حيث تخزين الذاكرة واستخدام وحدة المعالجة المركزية واستخدام النطاق الترددي. لذلك، استهدفنا في هذه الأطروحة تنفيذ الفئتين، وتحليلهما والمقارنة بين أدائهما لتحديد الآلية التي توفر الأداء الكفاء، ومن ثم اقتراح الحلول المناسبة لتحسين هذا الأداء.

من أجل تطبيق نظام ال VoWSN القائم على آلية الدفع الصوتي ، تم اقتراح وتنفيذ إطار لاستخدامه في معالجة الصوت وضغطه وتدقيقه. تم استخدام ثلاث تقنيات مختلفة لضغط البيانات الصوتية من أجل تقليل عرض الحزمة المطلوبة للنقل. ثم بعدها تم تقييم أداء الإطار باستخدام معلمات تقييم الأداء: تأخير أحادي الاتجاه ((One-Way Delay (OWD))، ارتعاش (jitter)، عرض الحزمة (B.W)، استخدام الذاكرة واستخدام وحدة المعالجة المركزية. علاوة على ذلك، تم تمديد العمل لتطبيق خدمة نقل الصوت عبر إنترنت الأشياء ((Voice over Internet of Things (VoIOTs)) عن طريق توصيل الإطار المقترح بالإنترنت من خلال شبكة Wi-Fi لدراسة وتحليل سلوك النظام عند الاتصال بالإنترنت. أظهرت النتائج أن أل OWD وال jitter يمكنهما أن يحافظا على قيود الزمن الحقيقي لشبكات VoWSN و VoIOTs مع ترددات أخذ عينات (sampling frequency) مختلفة.

من جهة أخرى، تواجه عملية أل ASR مع متطلباتها المعتادة تحديات خاصة في تصميم الخوارزميات والأجهزة عند تطبيقها باستخدام أجهزة أل WSN المحدودة الموارد. لذلك، هدفت الخطوة الثانية إلى تطبيق نظام أل ASR ذو الزمن الحقيقي والعالي الأداء على شبكة

أل WSN من خلال اقتراح آلية اتصال تستند على عملية تحسين نماذج القاموس (Dictionary) واللغة (Language Model) تلك المرتبطة بنظام أل ASR من خلال تحديدهما بحيز محدد من الجمل المطلوبة. تم اعتماد عمل النظام على بروتوكول مقترح يحول فئة النظام من نظام أل ASR المعتمد على الشبكة مع وجود قاموس ونموذج لغة كامل (أي مفردات كبيرة) إلى نظام أل ASR المضمن بالكامل مع نماذج مخصصة للقاموس واللغة (أي مفردات صغيرة). نتيجةً لذلك، فإن عملية التحول تلغي تبعية الشبكة التي تعد الميزة الرئيسية لمعظم أنظمة أداة التعرف على الكلام الحديثة. تم تقييم مقاييس الأداء في الزمن الحقيقي (معدل أخطاء الكلمة (WER))، وسرعة التمييز، والدقة، واستخدام وحدة المعالجة المركزية واستخدام الذاكرة) للنظام المقترح باستخدام نماذج القاموس واللغة المخصصة والكاملة. أظهرت نتائج تقييم الاداء تحسينا واضحا في أداء النظام المقترح عند تحسين نماذج القاموس واللغة حيث ان معدل أخطاء الكلمة (WER) قد قل بنسبة 67.11% وسرعة التمييز تحسنت بنسبة 3.96% والذاكرة المستخدمة قلت بنسبة 8.2%. بينما هدفت الخطوة الأخيرة إلى اقتراح إستراتيجية لتقييم أداء نظام أل ASR المقترح مقارنةً بنظام تدفق الصوت لتوضيح المكاسب (من حيث أداء الشبكة) التي تم الحصول عليها عن طريق تطبيق نظام أل ASR في نفس الظروف لكلا النظامين. تم تنفيذ هذه الخطوة من خلال نمذجة الأنظمة الأساسية المقترحة باستخدام حزمة محاكاة الشبكة أل OPNET. أظهرت النتائج التي تم الحصول عليها زيادة ملحوظة في سعة النظام عند استخدام نظام أل ASR المقترح مقارنة بنظام التدفق. أخيرًا، يتميز إطار ASR المقترح بكونه فعال من حيث التكلفة وسهل التثبيت وقابل للتوسعة والتكيف. بالنسبة لهذه الميزات، يمكن اعتباره حلاً بديلاً فعالاً لنظام تدفق الصوت لأنظمة الطوارئ المختلفة التي تعتمد على الاتصال الصوتي.

وزارة التعليم العالي والبحث العلمي

جامعة الموصل

كلية الهندسة



تحسين نظام نقل الصوت عبر شبكة المتحسس اللاسلكي

أطروحة مقدمة من قبل

إنعام فتحي خضر الأنعيمي

إلى

مجلس كلية الهندسة / جامعة الموصل كجزء من متطلبات

نيل شهادة الدكتوراه علوم في هندسة الحاسوب

هندسة الحاسوب/ شبكات الحاسوب

بإشراف

الأستاذ الدكتور

قتيبة إبراهيم علي

الأستاذ الدكتور

جاسم محمد عبد الجبار

2020 م

1441 هـ