

ระบบตอบรับโทรศัพท์อัตโนมัติ เพื่อสอบถามสภาพจราจร (Traffic Voice Information Service, TVIS)

02 565 7007

ณัฐนันท์ ทัดพิทักษ์กุล, ชัย วุฒิวิวัฒน์ชัย, อนันต์คณา ไซดิมงคล
หน่วยปฏิบัติการวิจัยวิทยาการมนุษยภาษา
ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

ในเวลาที่ประชาชนต้องการข้อมูลข่าวสารไม่ว่าจะเป็นสภาพจราจร รายงานอุบัติเหตุ รายงานการปิดถนน หรือแม้กระทั่ง รายงานสภาพอากาศ ต้องเป็นข้อมูลที่ทันต่อเหตุการณ์ เพราะต้องใช้ในการตัดสินใจเลือกเส้นทางในการเดินทาง ซึ่งในปัจจุบันประชาชนสามารถหาข้อมูลเหล่านี้ได้จากสื่อที่ให้บริการในรูปแบบต่างๆ โทรศัพท์ก็เป็นสื่อประเภทหนึ่งที่เข้าถึงได้ง่ายที่สุดไม่ว่าผู้โทรจะอยู่ที่ไหนแต่หากจะต้องใช้พนักงานมากคอยติดตามสภาพจราจร หรือสถานการณ์แล้วคอยตอบคำถามซ้ำๆ คงไม่ได้ประสิทธิภาพเท่าที่ควร

ศูนย์อิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ หรือ เนคเทค จึงได้มีการพัฒนาระบบ TVIS หรือ Traffic Voice Information Service ขึ้นให้เป็นระบบสอบถามข้อมูลจราจรในกรุงเทพฯ ที่หมายเลข 02-565-7007 โดยเป็น Automatic Call Center หรือ Interactive Voice Response (IVR) system ที่ผสานระหว่างการใช้เสียงพูดแบบอัตโนมัติ (Automatic Speech Recognition, ASR) ซึ่งช่วยให้คอมพิวเตอร์เข้าใจเสียงพูดของเราว่าเป็นคำว่าอะไร หรือเป็นการแปลง

เสียงพูดเป็นข้อความนั่นเอง เข้ากับระบบสังเคราะห์เสียงจากข้อความ (Text-to-Speech Synthesis, TTS) ซึ่งเป็นระบบที่สอนให้คอมพิวเตอร์อ่านข้อความภาษาไทย แล้วแปลงเสียงออกมาได้ ระบบ TVIS นี้จึงเป็น Automatic Call Center ที่ไม่ต้องใช้พนักงานในการรับโทรศัพท์

สำหรับขั้นตอนการทำงานของระบบ IVR เป็นดังรูปที่ 1 โดยจะเริ่มจากการรับเสียงพูดของผู้ใช้ผ่านโทรศัพท์ แล้วนำเสียงที่ได้มาแปลงเป็นข้อความ จากนั้นนำข้อความที่ได้ไปใช้ค้นหาคำตอบที่ผู้ใช้

ต้องการจากฐานข้อมูล ระบบจะตอบกลับเป็นเสียงพูดผ่านทางโทรศัพท์เช่นกัน โดยขั้นตอนทั้งหมดไม่ต้องอาศัย operator ที่เป็นมนุษย์แต่อย่างใด โดยข้อมูลจราจรที่ใช้ตอบคำถามจะดึงมาจาก www.traffy.in.th, www.trafficpolice.go.th และ www.traffic.thai.net โดยระบบนี้จะช่วยบริการข้อมูลจราจรแก่สังคมได้ตลอด 24 ชั่วโมง ช่วยลดต้นทุนการจ้างบุคลากร Call-center ช่วยประหยัดพลังงาน และประหยัดเวลาในการเดินทาง



รูปที่ 1 ระบบตอบรับอัตโนมัติสำหรับการรายงานสภาพจราจร (TVIS)

เทคโนโลยีที่ใช้ และเทคโนโลยีที่เกี่ยวข้อง

เทคโนโลยีที่ใช้ในระบบตอบรับอัตโนมัติมีอยู่ด้วยกัน 2 เทคโนโลยี คือ เทคโนโลยีการรู้จำเสียงพูด และเทคโนโลยีการสังเคราะห์เสียงพูดจากข้อความ ซึ่งเป็นเทคโนโลยีที่ทำให้คอมพิวเตอร์สามารถติดต่อสื่อสารกับคนด้วยวิธีการพูดได้

เทคโนโลยีการรู้จำเสียงพูดเรียกว่า การแปลงเสียงพูดเป็นข้อความ ช่วยให้ผู้ใช้สามารถติดต่อกับคอมพิวเตอร์ได้ด้วยเสียงพูด เทคโนโลยีการรู้จำเสียงพูดสามารถแบ่งกว้างๆ ตามรูปแบบการพูดออกได้เป็น 3 ระดับคือ

1. การรู้จำเสียงพูดคำโดด
2. การรู้จำเสียงพูดต่อเนื่องที่กำหนดโดยกรรมของประโยคพูด
3. การรู้จำเสียงพูดต่อเนื่องโดยไม่มีกำหนดโดยกรรมของประโยคพูด

หลักการรู้จำเสียงพูดอย่างง่ายคือการสกัดค่าทางด้านความถี่จากสัญญาณเสียงอินพุต นำไปเปรียบเทียบกับแบบจำลองเสียงของข้อความเอาท์พุตที่เป็นไปได้ทั้งหมด และเลือกเอาข้อความที่ให้ค่า

ความน่าจะเป็นสูงที่สุดเป็นผลลัพธ์ของการรู้จำ สำหรับ การรู้จำคำโดด ข้อความเอาท์พุตที่เป็นไปได้ประกอบด้วยคำโดดที่กำหนดไว้ในพจนานุกรมของระบบ แต่หากเป็นการรู้จำเสียงพูดต่อเนื่อง ข้อความเอาท์พุตที่เป็นไปได้จะเป็นประโยคที่ประกอบด้วยคำต่างๆ ที่กำหนดไว้ในพจนานุกรมของระบบ และเมื่อเป็นการรู้จำเสียงพูดต่อเนื่อง ความน่าจะเป็นของ

การเรียงคำเป็นประโยคจะมีความสำคัญต่อการตัดสินใจรู้จำ ค่าความน่าจะเป็นที่ได้จากการเทียบเสียงอินพุตกับแบบจำลองเสียงจะถูกถ่วงน้ำหนักเพิ่มความน่าจะเป็นที่ประโยคนั้นๆ จะเกิดขึ้นได้ในการใช้งานจริงด้วย

ระบบรู้จำเสียงพูดยังต้องคำนึงถึงจำนวนคำศัพท์ที่ระบบสามารถรู้จำได้ ระบบประยุกต์ที่สามารถสร้างได้เมื่อพิจารณารูปแบบของการพูดประกอบกับจำนวนคำศัพท์ แสดงดังรูปที่ 2

เทคโนโลยี ณ ปัจจุบัน สามารถประยุกต์ใช้การรู้จำเสียงพูดคำโดดและการรู้จำเสียงพูดต่อเนื่องที่กำหนดโดยกรรมของประโยคพูดในภาคธุรกิจได้แล้ว สำหรับการรู้จำเสียงพูดต่อเนื่องที่ไม่กำหนดโดยกรรมจะสามารถพัฒนาเป็นระบบที่ใช้งานได้เป็นที่พอใจของผู้ใช้ได้ด้วยเงื่อนไขสำคัญ อาทิเช่น สัญญาณเสียงพูดใช้งานระบบมีสัดส่วนสัญญาณเสียงต่อสัญญาณรบกวน (Signal to noise ratio, SNR) ไม่ต่ำจนเกินไป อาทิเช่น เกินกว่า 15 dB จำนวนคำศัพท์ที่ระบบต้องรู้จำได้ สอดคล้องกับ

การรู้จำเสียงพูดต่อเนื่องแบบไม่กำหนดโดยกรรม	กรอกหมายเลขรหัส	กรอกแบบฟอร์ม	ส่งข้อความด้วยเสียง	ตอบรับโทรศัพท์	สอบถามเส้นทาง	ถอดความการสนทนา	ถอดความรายงานข่าว		
การรู้จำเสียงพูดต่อเนื่องแบบกำหนดโดยกรรม	การรู้จำเสียงพูดคำโดด	ส่งข้อความอุปกรณ์	พูดชื่อต่อสาย	2	20	200	2000	20000	ไม่จำกัด

จำนวนคำศัพท์

รูปที่ 2: ระบบประยุกต์ที่ใช้เทคโนโลยีการรู้จำเสียงพูด เมื่อพิจารณารูปแบบการพูดและจำนวนคำศัพท์ที่จำได้



ขนาดของคลังข้อมูลเสียงและคลังข้อความที่สามารถนำมาใช้ในการฝึกฝนระบบรู้จำอาทิเช่น หากจำนวนคำศัพท์อยู่ในหลักไม่เกิน 3,000 คำ ควรมีคลังข้อความขนาดไม่ต่ำกว่า 1 ล้านคำในการฝึกฝนความน่าจะเป็นของการสร้างประโยค และใช้ข้อมูลเสียงไม่ต่ำกว่า 10 ชั่วโมง เพื่อการสร้างแบบจำลองเสียงพูด เป็นต้น

เทคโนโลยีสังเคราะห์เสียงพูดจากข้อความ เป็นเทคโนโลยีที่วิจัยและพัฒนาขึ้น เพื่อให้มนุษย์รับข้อมูลจากคอมพิวเตอร์ได้ด้วยการฟัง เสียงพูดที่ได้จากคอมพิวเตอร์ควรมีคุณภาพเสียงที่ใกล้เคียงเสียงพูดของมนุษย์ และมีความถูกต้องในการออกเสียงทั้งในแง่ความหมายหลักและความหมายแฝง และมีขั้นตอนการทำงานเป็นดังรูปที่ 3

ขั้นตอนการทำงานประกอบด้วย

1. ส่วนวิเคราะห์ข้อความ (Text analyzer) มีระบบแบ่งคำอัตโนมัติ (Automatic

word segmentation) ที่ได้รับการพัฒนาล่าสุด และระบบวิเคราะห์คำอ่าน (Letter-to-sound conversion) ซึ่งสามารถสร้างคำอ่านได้ครอบคลุมทุกคำในภาษาไทยเนื่องจากมีส่วนวิเคราะห์คำอ่านที่สามารถวิเคราะห์ได้แม้แต่คำที่ไม่ปรากฏในพจนานุกรม

2. ส่วนทำนายสัทสัมพันธ์ (Prosody prediction) ที่ช่วยในการวิเคราะห์ขอบเขตของวลีและการทำนายความยาวของหน่วยเสียง ทำให้เสียงสังเคราะห์มีความเป็นธรรมชาติดังเช่นเสียงพูดของคน
3. ส่วนสังเคราะห์เสียง (Speech synthesis) ได้ใช้เทคโนโลยีใหม่ล่าสุด ในการบีบอัดเสียงต้นแบบเก็บไว้ในแบบจำลองทางสถิติ ในขณะที่สังเคราะห์เสียง แบบจำลองทางสถิตินี้จะผลิตเสียงได้อย่างราบเรียบ ไม่เกิดการสะดุด

การประยุกต์ใช้กับนโยบาย

โดยปกติแล้วระบบ Call-Center ทั่วไปจะใช้ operator เป็นผู้ตอบคำถาม แต่ TVIS เป็นระบบตอบรับโทรศัพท์อัตโนมัติที่ใช้เทคโนโลยีการรู้จำเสียงพูดและการสังเคราะห์เสียงพูดจากข้อความ เป็นส่วนรับข้อมูลและตอบกลับผู้ใช้งานตามลำดับ ทำให้ TVIS เป็นระบบอัตโนมัติที่ใช้คอมพิวเตอร์พูดและตอบคำถามแทนคน

ขั้นตอนการทำงาน

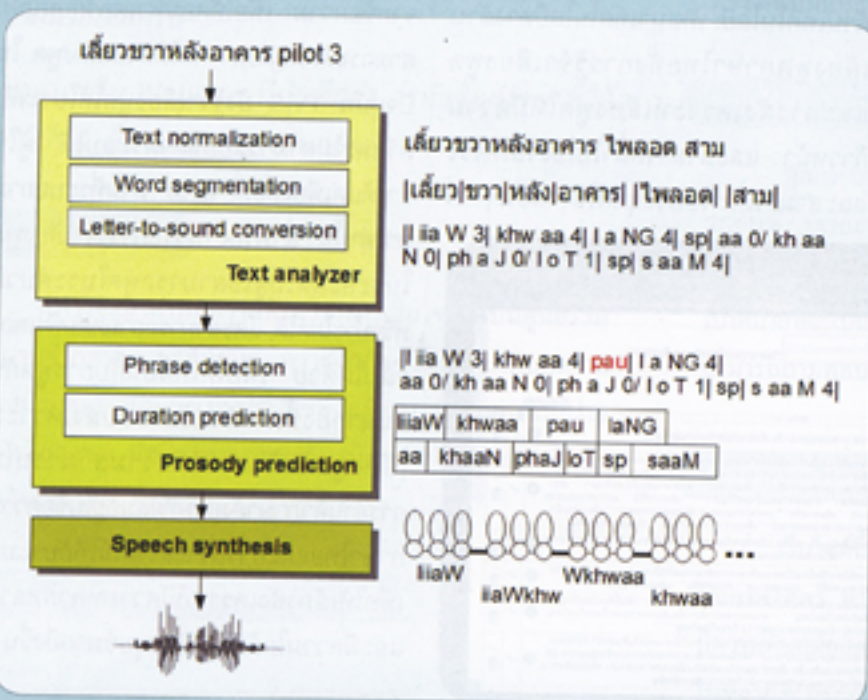
1. เมื่อผู้ใช้โทรศัพท์เข้ามาที่ 02-565-7007 ระบบจะให้พูดชื่อถนนที่ต้องการทราบสภาพจราจร เมื่อผู้ใช้พูดชื่อถนนเข้ามาในระบบจะแปลงจากเสียงพูดให้เป็นข้อความ

หมายเหตุ ระบบรู้จำเสียงพูดภาษาไทยที่ใช้ใน TVIS ในระยะแรกยังมีข้อจำกัดที่สามารถรู้จำเสียงพูดได้ในระดับคำเท่านั้น ทำให้ TVIS ในระยะแรกนี้ผู้ใช้งานสามารถพูดได้เฉพาะชื่อถนนในกรุงเทพเท่านั้น (รายชื่อถนนสามารถเข้ามาดูได้ที่ www.tvis.nectec.or.th ดังรูปที่ 4 หรือ เข้าไปฟังได้ที่ 02-565-7007 แล้วเข้าไปฟังที่เมนูรายชื่อถนน) เช่น จะถามสภาพจราจรถนนสีลมก็ให้พูดคำว่า "สีลม" เท่านั้นยังไม่สามารถพูดเป็นประโยคได้

2. ระบบจะนำชื่อถนนนั้นไปค้นในฐานข้อมูล ซึ่งฐานข้อมูลหลักที่ใช้คือ traffy.in.th ในขั้นตอนนี้จะมีการสรุปสภาพจราจรให้ได้สั้นที่สุด และข้อความที่ได้ต้องเป็นประโยคภาษาพูดด้วย

3. จากนั้นระบบจะนำข้อความที่ได้ไปผ่านระบบสังเคราะห์เสียงพูดจากข้อความ แล้วนำเสียงพูดที่ได้ไปเล่นให้ผู้ใช้งานฟัง

นอกจากนี้ TVIS ยังมีระบบให้พนักงานสามารถเข้ามาป้อนข้อมูลที่มีการตอบบ่อยๆ หรือข้อมูลที่ยังไม่มีการเอาขึ้น



รูปที่ 3: การสังเคราะห์เสียงพูดจากข้อความ



TVIS โทรฟรี 02-565-7007 สายด่วน 191 หรือ 112

จุดเริ่มต้น	ปลายทาง	ราคา	หมายเหตุ
กรุงเทพฯ	01	010	011
กรุงเทพฯ	02	020	021
กรุงเทพฯ	03	030	031
กรุงเทพฯ	04	040	041
กรุงเทพฯ	05	050	051
กรุงเทพฯ	06	060	061
กรุงเทพฯ	07	070	071
กรุงเทพฯ	08	080	081
กรุงเทพฯ	09	090	091
กรุงเทพฯ	10	100	101
กรุงเทพฯ	11	110	111
กรุงเทพฯ	12	120	121
กรุงเทพฯ	13	130	131
กรุงเทพฯ	14	140	141
กรุงเทพฯ	15	150	151
กรุงเทพฯ	16	160	161
กรุงเทพฯ	17	170	171
กรุงเทพฯ	18	180	181
กรุงเทพฯ	19	190	191
กรุงเทพฯ	20	200	201

รูปที่ 4 www.tvis.nectec.or.th

ระบบที่มีให้บริการ เช่น การรายงาน สถานการณ์การชุมนุม เส้นทางปิดถนน พร้อมบอกวิธีการหลีกเลี่ยง หรือแจ้งอุบัติเหตุ เป็นต้น โดยระบบนี้ข้อมูลจะเป็นดังรูปที่ 5 พนักงานป้อนข้อมูลเป็นข้อความเข้าระบบ จากนั้นระบบจะเปลี่ยนจากข้อความที่ได้ให้เป็นเสียงพูดด้วยระบบสังเคราะห์เสียงพูดจากข้อความ เห็นได้ว่าไม่ต้องเสียเวลาตามคนกดเสียงก็สามารถอัปเดตข้อมูลขึ้นระบบได้แล้ว ทำให้ข้อมูลของ TVIS มีข้อมูลที่มีความสดอยู่ตลอดเวลา

เวลา และเนื่องจาก TVIS ใช้คอมพิวเตอร์ในการรับสาย ทำให้เป็นการง่ายต่อการเพิ่มคู่สาย ช่วยลดต้นทุนการจ้างบุคลากร Call-center ช่วยประหยัดพลังงาน และประหยัดเวลาในการเดินทาง

ประโยชน์ของงานวิจัย

ด้านเทคโนโลยี: พัฒนาเทคโนโลยีทางด้านเสียงพูดภาษาไทยทั้งการรู้จำเสียงพูดและการสังเคราะห์เสียงพูดให้มีความก้าวหน้า และสามารถนำไปใช้งานได้จริง และสามารถนำไปประยุกต์ใช้งานอื่นๆ ที่

ไม่จำกัดแค่เพียงการรายงานสภาพจราจร เช่น รายงานเหตุการณ์การชุมนุม, บกจุดที่มีการปิดถนนและแนะนำเส้นทางเลี่ยงหรือเอาไปอ่านข่าวทั่วไปได้ เป็นต้น

ด้านสังคม: ช่วยบริการข้อมูลจราจรแก่สังคม ทำให้ผู้ขับที่มีข้อมูลสำหรับเลือกเส้นทาง ช่วยทำให้ประหยัดพลังงานและประหยัดเวลาการเดินทาง

ด้านเศรษฐกิจ: ช่วยลดต้นทุนการจ้างบุคลากรของ call-center และลดการนำเข้าอุปกรณ์ผ่านทาง

TVIS เป็นระบบบริการข้อมูลสภาพจราจร ผ่านระบบคอมพิวเตอร์อัตโนมัติเบอร์ 02-565-7007 ซึ่งเป็นอีกช่องทางหนึ่งให้ผู้ที่ต้องการเดินทางสามารถเข้ามาสอบถามสภาพจราจร เพื่อใช้เป็นข้อมูลในการตัดสินใจเลือกเส้นทาง

นอกจากนี้ TVIS เป็นตัวอย่างการประยุกต์ใช้งานเทคโนโลยีการรู้จำเสียงพูด และเทคโนโลยีการสังเคราะห์เสียงพูดจากข้อความ เพื่อให้คอมพิวเตอร์กับมนุษย์สามารถติดต่อสื่อสารกันได้ด้วยเสียงพูด ในปัจจุบัน TVIS ยังรู้จำเสียงพูดได้ในระดับคำและไม่สามารถอ่านตามระบบได้ ผู้ใช้ทำได้แค่เพียงตอบคำถาม ตามที่ระบบถามมาเท่านั้น สำหรับงานวิจัยการรู้จำเสียงพูดในระยะถัดไปผู้ใช้สามารถพูดในระดับวลีหรือประโยคได้ และสามารถถามคอมพิวเตอร์กลับได้ด้วย เพื่อให้เหมือนกับการพูดกับคนมากยิ่งขึ้น สำหรับระบบสังเคราะห์เสียงพูดจะมีการเพิ่มความสามารถในการอ่านคำภาษาอังกฤษที่ผสมอยู่ในข้อความภาษาไทยและการสร้างอารมณ์ให้กับระบบเพื่อให้เสียงสังเคราะห์มีความหลากหลายและมีความใกล้เคียงกับมนุษย์มากยิ่งขึ้น

TVIS โทรฟรี 02-565-7007 สายด่วน 191 หรือ 112

รูปที่ 5 ระบบการเพิ่มข้อมูลข่าวของระบบ TVIS

<http://www.nstda.or.th>