ในปัจจุบันมีการนำแบบจำลองฮิดเดนมาร์คอฟมาประยุกต์ใช้ในการรู้จำรูปแบบต่างๆ เป็นจำนวนมาก การรู้จำเสียงก็เป็นการรู้จำแบบหนึ่งที่นิยมใช้แบบจำลองฮิดเดนมาร์คอฟเป็นแบบจำลองทางเสียง การรู้จำด้วยเทคนิคนี้ประสิทธิภาพส่วนหนึ่งของการรู้จำเสียงจะขึ้นอยู่กับทอพอโลยีของแบบจำลองฮิดเดนมาร์คอฟที่เลือกใช้ ในวิทยานิพนธ์นี้นำเสนอวิธีการประมาณทอพอโลยีของแบบจำลองฮิดเดนมาร์คอฟสำหรับการรู้จำหน่วยเสียงในภาษาไทยโดยวิธีที่นำเสนอนั้นประกอบด้วย 2 ขั้นตอน ในขั้นตอนแรกเป็นการสร้างเซตของทอพอโลยีของแบบจำลองฮิดเดนมาร์คอฟด้วยการวิธีการประมาณทอพอโลยีที่เกิดจากการจับคู่กันระหว่างฟังก์ชันวัตถุประสงค์กับวิธีการผลิตทอพอโลยี ขั้นตอนต่อมาจะนำขั้นตอนวิธีเชิงพันธุกรรมมาประยุกต์ใช้เป็นขั้นตอนวิธีในการเลือกทอพอโลยีของแบบจำลองฮิดเดนมาร์คอฟที่มีความเหมาะสมสำหรับแต่หน่วยเสียงโดยคำนึงถึงความเหมาะสมโดยรวมและการเลือกทอพอโลยีนั้นจะเลือกจากเซตของทอพอโลยีที่สร้างขึ้นในขั้นตอนแรก จากผลการทดลองพบว่า ทอพอโลยีของแบบจำลองฮิดเดนมาร์คอฟที่ประมาณได้จากวิธีที่นำเสนอนั้นสามารถลดค่าความผิดพลาดในการรู้จำเสียงลงได้ 4.36% เมื่อเปรียบเทียบกับทอพอโลยีที่ใช้การเชื่อมต่อของสถานะแบบซ้ายขวา นอกจากนั้นทอพอโลยีดังกล่าวยังสามารถใช้งานได้ดี เมื่อสภาพแวดล้อมของเสียงที่ใช้เป็นฐานข้อมูลเพื่อประมาณทอพอโลยีแตกต่างจากสภาพแวดล้อมของเสียงในการรู้จำ

204502

The use of Hidden Markov Models (HMM) in many pattern recognition tasks is very common. Like other pattern recognitions, most Automatic Speech Recognition systems rely on HMM acoustic models. In such systems, recognition performances are significantly affected by their topologies. In this research, we proposed an HMM topology estimation approach for Thai phoneme recognition tasks whose process is divided into 2 stages. Firstly, a set of suitable topologies are constructed by combinations of different objective functions and topology generation methods. Second, a Genetic Algorithm is deployed as the topology selection algorithm which considers global fitness and selects the most suitable topology from the candidates proposed in the previous stage for each phoneme. As a result, the well-trained topology yields a maximum of 4.36% error reduction over predefined left-to-right models. The estimated topologies still work well when the topology estimation was performed on speech utterances whose recording environments differ from ones recognized.