Topic 4: Hidden Markov Model (HMM) for automatic speech recognition (ASR) with special name entities for dengue fever risky sites monitoring

Description: The objective of this project is to build a system of automatic speech recognition based on Hidden Markov Model (HMM) to recognize answers of users to the questions designed by the dengue fever risky sites monitoring of SWARM Project.

Risky sites for dengue transmission are places (mostly in urban areas) that are susceptible to make breeding sites. In particular, the presence of temporary open still water (in ponds, empty box, pots of flowers, etc.) favours mosquitoes breeding. For instance, pagodas and construction sites (for building houses, roads, bridges) have open still water and are therefore considered as risky sites.

Information about interest sites (like construction sites, pagodas, etc.) is gathered using speech with descriptive voices. The goal of this internship is to implement HTK Toolkit (from an existing library/framework) to perform automatic speech recognition of five site types: construction site, pagoda, rubbish, lake/pond, park. HTK Toolkit is HMM based approach that well applied for ASR purpose. With specific of name entities, for example of street names, pagoda names, construction names… from Hanoi areas, special language model will be applied for training to get higher accuracy for dengue fever risky sites monitoring purpose of SWARM project. General training data will be provided to train acoustic model plus recorded answering voices that will be collected based on questions designed by the dengue fever risky sites monitoring from SWARM Project.

Expected outcomes:
- Study, test, and implement a ASR system based on HTK Toolkit
- Build a speech database based on collected voices from the Topic1.
- User and technical documentations.

Supervisors:
- Luong Chi Mai, lcmai@ioit.ac.vn
- Nguyen Van Huy, huynguyen@tnut.edu.vn
- Bruno Lescalier, b.lescalier@posteo.de
- Nghiem Thi Phuong, nghiem-thi.phuong@usth.edu.vn

Pre-requisites:
- Motivations in research and innovation
- Good skills in programming
- Knowledge in signal, speech processing

References:

[1] Young, Steve. *The HTK Book*. UK: Cambridge University Engineering Department, 2009.

[2] Toolkit: http://htk.eng.cam.ac.uk/.

[3] Đặng Ngọc Đức, Nghiên cứu ứng dụng Mạng neuron và mô hình markov ẩn trong nhận dạng tiếng Việt, PhD Thesis, 2004.