

REZUMATUL TEZEI DE DOCTORAT

Contribuții la dezvoltarea sistemelor de verificare a vorbitorului pe echipamente mobile

Autor: ing. Florin-Dragoș CURELARU, *e-mail:* patronixys@yahoo.com

Conducător de doctorat: Gl. bg. (rtg) prof. univ. dr. ing. Mihai RADU

Scopul principal al tezei a fost găsirea celor mai robuste metode de verificare independente de text a vorbitorului și a parametrilor optimi ai metodelor, care să conducă la cele mai bune performanțe, în condițiile antrenării pe cantități limitate de date de la clienți și evaluării pe rostiri scurte colectate în condiții reale. Pentru ca performanțele metodelor de verificare a vorbitorului să fie evaluate în condiții reale, s-au efectuat experimente cu rostiri colectate pe echipamente mobile, care să implice diverse medii acustice și variate tipuri de microfoane. Cantitățile limitate de date de antrenare colectate de la clienți și lungimea scurtă a rostirilor de evaluare sunt motivate de două caracteristici dezirabile ale unui sistem de verificare: să înroleze rapid noi clienți și să nu ceară clienților un timp considerabil pentru a se autentifica.

Experimentele din teză s-au efectuat cu rostiri din două corpuri disponibile publicului în mod gratuit: *MIT Mobile Device Speaker Verification Corpus* (MIT-MDSVC) și *CSTR Voice Cloning Toolkit Overview* (CSTR VCTK). Corpusul MIT-MDSVC conține o cantitate limitată de rostiri înregistrate pe echipamente mobile în trei medii diferite, cu două microfoane. Corpusul CSTR VCTK este mult mai voluminos ca MIT-MDSVC și a fost utilizat, în special, pentru a urmări evoluția performanței sistemelor de verificare a vorbitorului la extinderea lotului de antrenare cu rostiri de la vorbitori neînrolați, colectate după alte criterii privind mediile acustice ale înregistrărilor.

Având în vedere modelarea folosită la reprezentarea vorbitorilor și schema utilizată pentru clasificarea vorbitorilor, abordările moderne de verificare a vorbitorilor aparțin în general uneia din cele două categorii: generativă sau discriminativă. Cum fiecare categorie are avantaje specifice care nu se regăsesc în cealaltă categorie, în încercarea de a beneficia de avantajele ambelor categorii a apărut o categorie hibridă de metode construite pe modele generative antrenate în mod discriminativ. Pornind de la clasificarea metodelor de verificare în generative, discriminative și metode construite pe modele generative antrenate discriminativ, s-au ales pentru evaluare comparativă cinci metode reprezentative, cel puțin una din fiecare categorie. Cele cinci metode reprezentative sunt: *Gaussian Mixture Model – Universal Background Model* (GMM-UBM), reprezentarea rostirilor prin vectori identitate în conjuncție cu *Analiza Discriminant Liniară Probabilistă* (din categoria metodelor generative), mașini cu vectori suport (engl. "Support Vector Machines" - SVM) cu nucleu polinomial (din categoria metodelor pur discriminative), și pe SVM cu nucleu Fisher și cu nucleu liniar GMM Supervector (din categoria metodelor construite pe modele generative antrenate discriminativ).

Pentru o evaluare cât mai obiectivă a metodelor de verificare s-a efectuat un set considerabil de experimente, care să evidențieze influența detectoarelor de voce în performanțele sistemelor de verificare. S-au ales patru detectoare de voce populare: un detector de voce bazat pe măsurarea energiei inclus în setul de instrumente *Alize/LIA_RAL*, un detector de voce autoadaptiv propus de Kinnunen et al., un detector de voce statistic propus de Sohn et al. și detectorul de voce standardizat G.729B.