

Kurzfassung der Studienarbeit

Abteilung	Informatik
Namen der Studierenden	Reto Brühwiler Manuela Grob
Studienjahr	HS 2009/2010
Titel der Studienarbeit	ATM-Voice
Examinator	Prof. Dr. P. Heinzmann
Themengebiet	Internet-Technologien und -Anwendungen
Projektpartner	Siemens Schweiz
Institut	ITA: Institut für Internet-Technologien – und Anwendungen

Im Rahmen der Studienarbeit Air-Traffic-Monitoring (ATM) – Voice sollte abgeklärt werden, wie gut sich die Inhalte von Flugfunkmeldungen erkennen lassen. Insbesondere ob sich der Funkverkehr zwischen dem Tower und dem Flugzeug automatisch in Textmeldungen umwandeln liesse. Schlussendlich sollten diese Meldungen einer Flugnummer zugeordnet werden können.

Im ersten Teil der Arbeit wurden die Grundlagen der Spracherkennung erarbeitet. Dazu gehört das Verständnis der einzelnen Komponenten, wie zum Beispiel das Hidden Markov Model (HMM) und Begriffe wie Mel frequency cepstral coefficients (MFCC) und Perceptual Linear Prediction (PLP). In der Einarbeitungsphase musste man sich auch mit den Flugfunkregeln vertraut machen. Es zeigte sich, dass Flugfunkmeldungen auf einem kleinen Vokabular und auf einer einfachen Grammatik basieren, was die Spracherkennung eigentlich erleichtern würde. Dem gegenüber wirken sich die schlechte Tonqualität und die sehr unterschiedliche Sprecherqualität erschwerend auf die Spracherkennung aus.

Unter Berücksichtigung dieser Erkenntnisse wurde die Anforderungsanalyse gemacht. Darauf basierend wurden verschiedene Speech-to-Text Lösungen im Hinblick auf den Einsatz im Flugfunkbereich gesucht und verglichen. Für die Realisierung eines Prototypen-Systems wurde das Open Source Framework Sphinx 4.0 ausgewählt. Ausschlaggebend waren neben den Kosten (Freeware) und der Verfügbarkeit vor allem die Anpassbarkeit der einzelnen Komponenten. Bei Sphinx können gemäss Dokumentation eigene „Dictionaries“ und „Grammatiken“ vorgegeben werden. Somit hätte ein „Dictionary“ erstellt werden können, welches speziell an den Wortschatz der Fliegerei angepasst ist und eine „Grammatik“ anhand der Flugfunkregeln ermöglicht.

Bei der Realisierungsphase hat sich aber leider gezeigt, dass sich die Sphinx Komponenten nicht wie versprochen anpassen liessen. Probleme bereitete insbesondere die Kombination des geeignetsten AcousticModels und der eigenen Grammatik.

Um die Funksprüche soweit in Text umwandeln zu können, wie das eigentliche Ziel wäre, braucht es noch einiges an Forschungs- und Entwicklungsarbeit. Nichtsdestotrotz konnte mit dieser Arbeit eine gute Basis geschaffen werden. Vor allem in der Erstellung des Dictionaries und der Grammatik, welche zu einem grossen Teil ausgearbeitet wurde.