Towards Ambient Search ⋆

Stephan Radeck-Arneth[1,2], Chris Biemann[2], and Dirk Schnelle-Walka[1]

[1] Telecooperation, Dep. Computer Science, TU Darmstadt, Germany
[2] Language Technology, Dep. Computer Science, TU Darmstadt, Germany

Abstract. In ongoing discussions participants tend to pick up their smart phones to retrieve relevant information for clarification, severely hampering the flow of the discussion. We introduce ambient search as a variant of information retrieval where a system unobtrusively provides relevant information snippets in the background without the need to steer devices actively. In this demo paper, we describe a first prototype of our ongoing research activities towards such a system.

1 Introduction

Phubbing describes a social problem where others are being ignored in favor of a mobile phone [3]. This may be done to retrieve information for clarification of facts for the current discussion, but still hampers the flow of conversation. We propose a system, which follows the discussion and returns related information in the background without requiring users to pick up and actively interact with devices. We define this as *ambient search*, featuring (i) real time information retrieval, (ii) presentation of topic-related information snippets and (iii) passive behavior. The system will unobtrusively present topic-related information snippets while the discussion continues. The dialog partners may or may not use them to collaboratively retrieve more detailed information.

We place ambient search into the continuum between Human Computer Interaction and Information Retrieval (IR). In this paper we introduce our first efforts towards the realization of such a system.

2 Related Work

Anzalone et al. [1] introduced a topic recognition system for social robots. They define TF-ITF to calculate the relevance per word to predefined topics. The definition relates to TF-IDF that describes the relevance of words to a document. Similarly, words with a high TF-ITF weight are considered to be more relevant for a topic. They also consider topic recognition to be helpful in the presence of speech recognition errors. Along these lines, Stas et al. [8] suggested an algorithm to build robust language models for a specific domain. They separated heterogeneous text data into binary domain classes that

 In: T. Seidl, M. Hassani, C. Beecks (Eds.): Proceedings of the LWA 2014 Workshops: KDML, IR, FGWM, Aachen, Germany, 8-10 September 2014, published at http://ceur-ws.org

improved the model perplexity for a transcription and dictation system for the Slovak language. Such a strategy might be helpful in detecting topics and improve recognition accuracy.

Snippets are a well-known strategy for text summarization [6] and feature characteristics of dynamic summaries. Consequently, their computation needs to be fast since they cannot be precomputed off-line and must be synthesized based on the query results.

Personalized IR systems are, amongst others, investigated by Jeh et al. [4]. They extended the well-known PageRank algorithm to a personalized form.

3 Approach

An overview of the envisioned architecture is shown in Fig. 1. We regard the interplay of components as information streams.

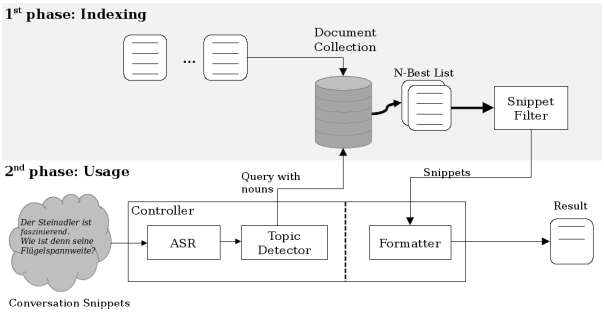

Fig. 1. Ambient Search Architecture

An *automated speech recognizer* (ASR) continuously processes the audio input stream of an ongoing discussion and forwards the recognized utterances to a *topic detector* to extract the meaningful parts. For now, we restrict this to detection of nouns. We are currently investigating the appropriateness of the selection process and strategies to combine the most recent nouns in the streams into subsets with various logical operators for an optimized balance between precision and coverage. Moreover, this module queries a structured document collection for a set of related documents. A *snippet filter* is responsible for filtering these documents to snippets, i.e. the relevant passages within a document. A *formatter* then highlights the topic identifiers (nouns in our case) to make the appearance more comprehensible. Hence, we expect users to be able to easily understand the causal relation between speech and the presented results. At this stage of development, we do not consider possible problems as a result of context switches during discussions or cross-talk.

We see the following advantages of this approach: (i) The assistant stays in the background without disturbing the user. (ii) The user can access the displayed snippets on demand. (iii) The snippet continuously updates the available snippets. However, these advantages will have to be validated in user studies.

Our current prototype provides basic implementations for all needed components and enables us to get first experiences with ambient search. We employed Sphinx [5] for ASR using our own models for German [7]. Nouns are identified by a pretree-based POS-Tagger [2]. For the document collection we are using the German Wikipedia indexed by Solr[3].

[3] http://lucene.apache.org/solr/

4 Conclusion & Future Work

We took the first steps towards our vision of ambient search as a basis for further investigation. Future and ongoing research activities cope with extracting appropriate keywords from the dialog stream. For this, we are currently inspecting transcribed dialogs to find possible strategies and we are developing an automated evaluation framework. Furthermore we are looking into user interaction with ambient search and its applicability to group discussions. Another important task we have to tackle for making ambient search practicable is to improve ASR performance, especially in noisy environments.

Acknowledgements

This work was partly supported by the Bundesministerium für Bildung und Forschung (BMBF), Germany under the programme "KMU-innovativ: Mensch-Technik-Interaktion für den demografischen Wandel".

References

1. Anzalone, S.M., Yoshikawa, Y., Ishiguro, H., Menegatti, E., Enrico, P., Sorbello, R.: A topic recognition system for real world human-robot conversations. In: Intelligent Autonomous Systems, vol. 12, pp. 383–391. Springer (2013)
2. Biemann, C., Quasthoff, U., Heyer, G., Holz, F.: ASV Toolbox: a Modular Collection of Language Exploration Tools. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08). pp. 1760–1767. European Language Resources Association (ELRA), Marrakech, Morocco (2008)
3. Coehoorn, M.: Phubbing? An absurd design intervention for redefining smart-phone usage. Master's thesis, TU Delft, Delft University of Technology (2014)
4. Jeh, G., Widom, J.: Scaling personalized web search. In: Proceedings of the 12th International Conference on World Wide Web. pp. 271–279. WWW '03, ACM (2003), http://doi.acm.org/10.1145/775152.775191
5. Lamere, P., Kwok, P., Gouvêa, E., Raj, B., Singh, R., Walker, W., Warmuth, M., Wolf, P.: The CMU Sphinx-4 Speech Recognition System. In: IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2003), Hong Kong. pp. 2–5 (2003), http://www.cs.cmu.edu/~rsingh/homepage/papers/icassp03-sphinx4_2.pdf
6. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to information retrieval, vol. 1. Cambridge University Press Cambridge (2008)
7. Schnelle-Walka, D., Radeck-Arneth, S., Biemann, C., Radomski, S.: An Open Source Corpus and Recording Software for Distant Speech Recognition with the Microsoft Kinect. In: Speech Communication; 11. ITG Symposium. p. 4. VDE (2014), (to appear)
8. Stas, J., Juhar, J., Hladek, D.: Classification of heterogeneous text data for robust domain-specific language modeling. EURASIP Journal on Audio, Speech, and Music Processing 22 (2014), http://dx.doi.org/10.1186/1687-4722-2014-14