Web-based Machine Translation

Haifeng Wang
Baidu
Beijing, 100085, China
`wanghaifeng@baidu.com`

1 Abstract

Machine translation (MT) has been studied for more than 60 years. World-Wide-Web offers more opportunities to MT. We could try to crawl more web data to train the MT system. But we have to filter the very noisy web data. There are many potential web-based applications for MT, such as translation of web-page, translation of instant message, translation of SNS, translation of e-commerce, mobile translation, etc. To make better use of the web data, and to produce better web-based MT applications, we should also adapt the MT methods to the web scenario. In this talk, I will introduce our work on web-based machine translation.

2 Biography

Dr. WANG Haifeng a senior scientist at Baidu, and a visiting professor at Harbin Institute of Technology. At Baidu, he is the head of Baidu's NLP department, and the advisor of its speech team, the technical leader of its recommendation & personalization team, and one of the core members of Baidu's technology committee. He received his PhD in computer science from Harbin Institute of Technology in 1999. He worked as an associate researcher at Microsoft Research China 1999 2000, a research scientist at iSilk.com (Hong Kong) 20002002, and chief research scientist and deputy director at Toshiba (China) R&D Center till Jan. 2010. He has authored more than 70 NLP papers, including 13 full papers in ACL main conferences. His research interests span a wide range of topics including: MT (SMT, RBMT, EBMT, TM and hybrid methods), parsing, generation, grammar induction, paraphrase, collocation extraction, SRL, WSD, LM, recommendation, personalization, speech and search. He has served as program chair, area chair, tutorial chair, workshop chair, industry track chair and PC members for numerous NLP conferences including ACL, SIGIR, NAACL, EMNLP, COLING and IJCNLP, etc. He also serves as associate editor of ACM TALIP, guest editor of ACM TIST. He is the Vice-President-Elect of the ACL.