

Nokia-Tsinghua Academia Summit



Short Message: Benefits and Challenges

Guo Jun

School of Information Engineering, BUPT

Unexpected Popularity

- ∅ SMS(Short Message Service) is presented only as an added value service, but it gets great popularity in eastern countries like China, Japan and Korea.
- ∅ In China, almost every user sends SMs in everyday, whereas little part of them call in everyday.
- ∅ SMS has become one of the most important incomes of mobile operators
 - ∅ 12 bln SMs were sent in the 7days of this year's Spring Festival.



The Transmission Features of SMS

- ✓ The SMS has distinguished transmission features:
 - ✓ Higher confidentiality than call
 - ✓ Higher rapidity than email
 - ✓ Higher reachability than call and email
 - ✓ Higher diffusibility than call and email, and
 - ✓

More and More Useful Functions

- ✓ The functions of SMS become more and more.
 - ✓ Daily private communications
 - ✓ Greetings in festivals and great days
 - ✓ Sharing jokes and humors with friends
 - ✓ Broadcasting news and announcements
 - ✓ Delivering literary works like novels, essays etc.
 - ✓

More and More Illegal Utilizing

- ✓ The features of transferring info by SMs are also favored by illegal people, through SMS, they
 - ✓ spread messages to goad violence,
 - ✓ send legally forbidden contents,
 - ✓ sell illegal products,
 - ✓ make traps to bilk ordinary people, and
 - ✓

The Challenges We are Facing

- ✓ To block the illegal messages without hurting the privacies of ordinary people.
- ✓ To provide more convenient and more rapid means to input SMs.
- ✓ To reduce the SM's minus effects to the regular languages.
- ✓



The Key Scientific Problems

- ✓ SM's language features and their modeling
- ✓ Category Systems
- ✓ Classification and clustering algorithms
- ✓ Spam filtering algorithms
- ✓ Multimodality input of SMs, and
- ✓

The Linguistic Features of SMs

- ✓ Short length
- ✓ Irregular grammar
- ✓ Non punctuation
- ✓ Novel word usage
- ✓ Homophones and input errors, and
- ✓

The Difficulties in Present Methods

- ✓ The sparseness of data in VSM and n-gram Model
- ✓ The plenty of noise in data
- ✓ Special rules for word segmentation
- ✓ Special term frequencies
- ✓ Special transfer probabilities between terms, and
- ✓

Current Researches in China

- ✓ Funded projects
 - ✓ The classification and filtering on SMs (NSFC)
 - ✓ Information monitoring in SMS (MII)
 - ✓ Spam filtering (China Mobile, etc)
 - ✓ SMS' impacts on culture and literature (CASS)
- ✓ Research units
 - ✓ Universities: BUPT, BIT, HIT
 - ✓ Institutes: ICT of CAS, CASS
 - ✓ Enterprises: China Mobile, China Unicom,

Nokia-Tsinghua Academia Summit



Thanks !