Combination of DTW-based and CRF-based Spoken Term Detection on the NTCIR-11 SpokenQuery\&Doc SQ-STD Subtask
Hiromitsu Nishizaki, Naoki Sawada, Satoshi Natori (University of Yamanashi)
Kentaro Domoto, Takehito Utsuro (University of Tsukuba)

1 Introduction
- We proposed a two-stage STD method
- Combination of Dynamic Time Warping (DTW) based and CRF-based methods
- The DTW-based method outputs so many false detections
- Goal of this research:
  - To control false detections by the DTW method using CRF-based triphone detection framework

2 Two-pass STD framework

Overview of two-pass STD system

1st pass (DTW approach)
- Query term
- Converting to phoneme sequence
- DTW-based term search engine
- Calculation of term detection probability
- CRF models
- Re-ranking
- STD Result

2nd pass (CRF approach)
- PTN-formed index

PTN-formed index by 10 ASR systems

<table>
<thead>
<tr>
<th>ASR ID</th>
<th>Outputs of 10 ASRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASR #1</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #2</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #3</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #4</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #5</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #6</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #7</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #8</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #9</td>
<td>f u j i s a N</td>
</tr>
<tr>
<td>ASR #10</td>
<td>f u j i s a N</td>
</tr>
</tbody>
</table>

Training example of “n-e-p” detection model by CRF
All triphones are modeled by this framework.

3 STD result

Evaluation data:
NTCIR-11 SpokenQuery\&Doc SQ-STD subtask
CRF model training data:
CORE lecture set of Corpus of Spontaneous Japanese

ROC curve

CRF-based term detection

CRF-based detection approaches

4 Conclusion
- We proposed the two-stage STD method using the combination of DTW-based and CRF-based detection approaches
- The CRF-based STD approach did not improve the DTW-based STD results
- Because the CRF models were trained from the CSJ speeches (mismatch)