CMNEE: A Large-Scale Document-Level Event Extraction Dataset based on Open-Source Chinese Military News

Mengna Zhu, Zijie Xu, Kaisheng Zeng, Kaiming Xiao, Mao Wang, Wenjun Ke, Hongbin Huang

LREC-COLING 2024

Background

- Events are the basic units of human activities and interactions, containing rich information.
- Event extraction refers to the extraction of structured information from unstructured text, which is typically separated into two subtasks: event detection and event argument extraction.
- Current research mainly focuses on the general news (ACE 2005, MAVEN, etc) or financial domains (ChfinAnn, Duee-fin), with only relatively fewer studies for military domain, which impedes the study of event extraction in this domain.

Background

Existing datasets for EE related tasks

| Level | Dataset | Domain | Language | Docs | Evtypes | ArgRs | Events | ED | EAE |
|-------|------------|-----------|----------|--------|---------|-------|---------|--------------|--------------|
| Sent | ACE 2005 | general | English | 599 | 33 | 35 | 4,090 | \checkmark | \checkmark |
| | MAVEN | general | English | 4,480 | 168 | - | 111,611 | \checkmark | × |
| | DuEE | financial | Chinese | 11,224 | 65 | 121 | 19,640 | \checkmark | \checkmark |
| | MNEE | military | Chinese | 13,000 | 8 | 10 | 6,997 | \checkmark | \checkmark |
| Doc | RAMS | general | English | 3,993 | 139 | 65 | 9,124 | \checkmark | \checkmark |
| | WikiEvents | general | English | 246 | 50 | 59 | 3951 | \checkmark | \checkmark |
| | Duee-fin | financial | Chinese | 11,700 | 13 | 92 | 11,031 | \checkmark | \checkmark |
| | ChfinAnn | financial | Chinese | 32,040 | 5 | 24 | 47,824 | \checkmark | \checkmark |
| | DocEE | general | English | 27,485 | 59 | 356 | 27,485 | × | \checkmark |
| | CMNEE | military | Chinese | 17,000 | 8 | 11 | 29223 | \checkmark | \checkmark |

Background

An example of CMNEE

S01: In this air campaign to attack the Taliban and Osama bin Laden in Afghanistan, the UK has employed two nuclear-powered attack submarines.
S02: The two submarines were the British Navy Trafalgar class, HMS Trafalgar and HMS Triumph.
S03: One of them, HMS Trafalgar, is a British submarine that participated in the Argonaut 2001 exercise and is permanently deployed in the Mediterranean Sea

S06: The first firing of Tomahawk cruise missiles from a British naval vessel took place in 1999, when the British Fast Class submarine HMS Splendor fired multiple missiles attacking targets in Yugoslavia.
S07: Originally, the UK designed the Trafalgar class submarine for the North Atlantic and Mediterranean confrontation with the former Soviet Union during the Cold War.

| Conflict (attack) | | | | |
|-------------------|-----------------|--|--|--|
| Subject | UK | | | |
| Object | Osama bin Laden | | | |
| Location | Afghanistan | | | |

| Conflict (attack) | | | | |
|-------------------|-------------|--|--|--|
| Subject | UK | | | |
| Object | Taliban | | | |
| Location | Afghanistan | | | |

| Deploy (deployed) | | | | |
|-------------------|-------------------|--|--|--|
| Subject | British Navy | | | |
| Militaryforce | HMS Trafalgar | | | |
| Location | Mediterranean Sea | | | |

| | UK, British, UK | | | |
|---------------------------|---------------------------------|--|--|--|
| Co-reference arguments | British Navy, British naval | | | |
| 3.92 | HMS Trafalgar, HMS Trafalgar | | | |

| Conflict (attacking) | | | | |
|----------------------|-----------------------|--|--|--|
| Subject | British naval | | | |
| Object | targets in Yugoslavia | | | |
| Date | 1999 | | | |

| Conflict (co | Conflict (confrontation) | | | | | |
|--------------|--------------------------|--|--|--|--|--|
| Subject | UK | | | | | |
| Object | the former Soviet Union | | | | | |
| Location | the North Atlantic | | | | | |
| Date | the Cold War | | | | | |

| Conflict (confrontation) | | | | | |
|--------------------------|-------------------------|--|--|--|--|
| Subject | UK | | | | |
| Object | the former Soviet Union | | | | |
| Location | Mediterranean | | | | |
| Date | the Cold War | | | | |

Construction

Methodology



Construction

Event schema

| Event type | | Argum | ent role | |
|------------|---------|---------------|----------|-----------|
| Experiment | Subject | Equipment | Date | Location |
| Manoeuvre | Subject | Content | Date | Area |
| Deploy | Subject | Militaryforce | Date | Location |
| Support | Subject | Object | Date | Materials |
| Accident | Subject | Result | Date | Location |
| Exhibit | Subject | Equipment | Date | Location |
| Conflict | Subject | Object | Date | Location |
| Injure | Subject | Quantity | Date | Location |

Analysis







Figure 1 Data sources

Figure 2 Event type distribution

Figure 3 Multi-events distribution

Overlapping events proportion: 42% Long arguemnts (more than 10 characteristics): 17%

Evaluation

Overall results

| Models | Event Detection | | | Event Argument Extraction | | |
|------------|-----------------|------|------|---------------------------|------|------|
| | Р | R | F1 | Р | R | F1 |
| DCFEE-O | - | - | - | 30.3 | 22.3 | 25.7 |
| DCFEE-M | - | - | - | 26.4 | 22.0 | 24.0 |
| GreedyDec | - | - | - | 39.4 | 19.9 | 26.4 |
| Doc2EDAG | - | - | - | 54.3 | 23.9 | 33.2 |
| DEPPN | - | - | - | 38.2 | 35.0 | 36.5 |
| BERT+CRF | 73.1 | 77.7 | 75.3 | 63.1 | 52.3 | 57.2 |
| EEQA | 65.8 | 80.5 | 72.4 | 39.0 | 39.1 | 39.0 |
| TEXT2EVENT | 30.1 | 60.6 | 40.2 | 31.3 | 41.3 | 35.5 |
| PAIE | - | - | - | 72.0 | 67.0 | 69.4 |

Evaluation





Error analysis and expected directions

- Identification Mistakes (nearly 50%)
 - models can better understand event semantic information based on complex text
- Majority Bias (about 20%)
 - models can cope with data imbalance and accomplish the extraction of fewer sample events more efficiently
- Extraction Boundary (more than 30%)
 - models can better determine the extraction boundaries so that the extracted information is concise and effective

Thanks!