

# Factionalism and the Red Guards under Mao's China: Ideal Point Estimation Using Text Data

David, Yen-Chieh Liao<sup>1</sup>   Yi-Nung Tsai<sup>3</sup>   Daniel Tene<sup>2</sup>  
Dechun Zhang<sup>4</sup>

<sup>1</sup>PhD Candidate at the University of Essex and Pre-Doctoral Researcher at  
University of Bamberg

<sup>2</sup>PhD Student, the University of Essex

<sup>3</sup>PhD in Political Science, National Chengchi University

<sup>4</sup>PhD Candidate, Institute for Area Studies of Faculty of Humanities, Leiden  
University

8th May 2022

## Motivation

- we design a new strain of text scaling method, Swordfish (Slogan-featured Wordfish), that takes advantage of the TextRank algorithm to extract the most representative political slogans in a given context and estimates Wordfish with those extracted text variables.
- We test this method using the case of the Great Proletarian Cultural Revolution in China and the historical archive from 1966.
- The results estimated by our approach are shown to be consistent with the representative qualitative literature of factionalism regarding the Cultural Revolution.

## Difficulties in Generic *Wordfish* Model

- we cannot directly perceive whether we have enough unique words across each document. In general, except for a few words associated with political preference, most terms are not as straightforward as to be able to identify their political concepts, particularly in Chinese and Japanese.
- Critics of *Wordfish* are quick to point out that such word frequency-based approaches is ignorant of semantic structure and underlying context (p.325 S.-O. Proksch and Slapin 2009).
- Difficulties in implanting *wordfish* in no space language such Chinese and Japanese

# Background

## ■ The Most Representative Major Organizations and Participants of the Red Guards

Factions	Student Participants	Characteristics
The Conservatives	Tsinghua University High School Red Guards (清华附中红卫兵), Western District Picket Corps (西纠), The United Action Committee (联动)	<ul style="list-style-type: none"><li>■ High school students;</li><li>■ Participant whose family worked in the bureaucracy system ;</li><li>■ Support to Liu Shaoqi, Deng Xiaoping, and party (and government) officials</li></ul>
The Rebels	Nie Yuanzi (聂元梓), Kuai Dafu (蒯大富), Third Headquarters (三司), Tsinghua Jinggangshan Regiment (清华大学井冈山兵团), Capital Red Guards Congress (首都大专院校红卫兵代表大会)	<ul style="list-style-type: none"><li>■ College students;</li><li>■ Participants from colleges and were active at the late stage of the movement;</li><li>■ The criticism of Liu Shaoqi, Deng Xiaoping, and party (and current government) officials</li></ul>

# Data



(a) Students writing big-character posters at Beijing Normal University



(b) Big-character posters during China's Cultural Revolution

Figure: Big-character posters during China's Cultural Revolution

# Data

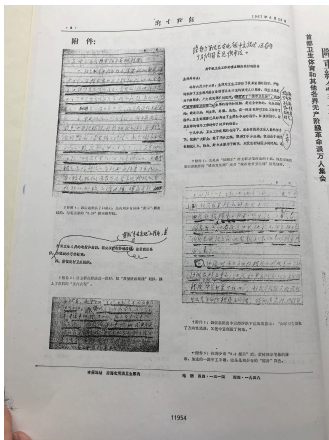


Figure: Self-printed tabloids (Xiaobao 小报) released by Red Guard students.  
Source: Center for Chinese Research Materials (2001)

David, Yen-Chieh Liao, Yi-Nung Tsai, Daniel Tene, and Dechun Zhang  
Factionalism and the Red Guards under Mao's China

# The Keyword Extraction Scaling Approach

## ■ Automated Keyword Extraction - TextRank

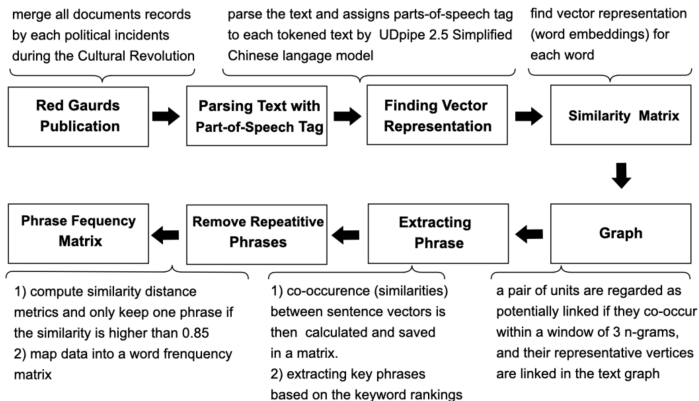
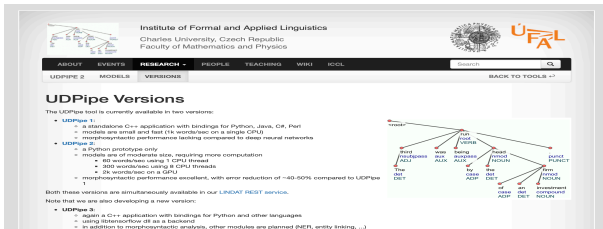


Figure: The flowchart of TextRank for the text pre-processing task

# The Keyword Extraction Scaling Approach

## UDpipe 2.5 Language Model

- we apply a pre-trained model from UDpipe, with **parts of speech (POS) tagging** and **linguistic dependencies**, to tokenize the collected textual data (Straka, Hajič, and Straková 2016; Straka and Straková 2017), see <https://ufal.mff.cuni.cz/udpipe/2>.



The screenshot shows the UDpipe website interface. At the top, it features the logo of the Institute of Formal and Applied Linguistics (UFAL) and Charles University, Czech Republic. Below this is a navigation bar with links for ABOUT, EVENTS, RESEARCH, PEOPLE, TEACHING, WIKI, and KCCL. A search bar is located on the right. The main content area is titled "UDPipe Versions" and lists two versions:

- UDPipe 1:**
  - is a standalone C++ application with bindings for Python, Java, C#, Perl
  - models are small and fast (1k words/sec on a single CPU)
  - morpho-syntactic performance lacking compared to deep neural networks
- UDPipe 2:**
  - is a Python prototype only
  - models are of moderate size, requiring more computation
    - ~60 words/sec using 1 CPU thread
    - ~300 words/sec using 8 CPU threads
  - ~2k words/sec on 8 GPUs
  - morpho-syntactic performance excellent, with error reduction of ~40-50% compared to UDpipe 1

Below the version list, it states that both versions are simultaneously available in their **LINGUAT REST service**. A note mentions that a new version is also being developed:

- UDPipe 3:**
  - again a C++ application with bindings for Python and other languages
  - using tensorflow dl as a backend
  - in addition to morpho-syntactic analysis, other modules are planned (NER, entity linking, ...)

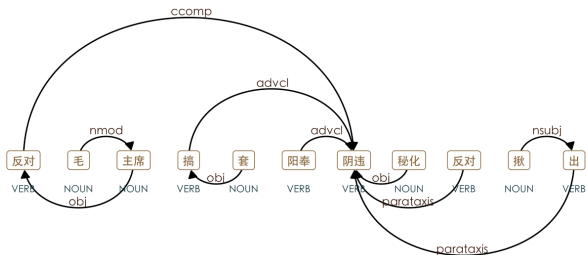
On the right side of the page, there is a dependency parse tree for the sentence "The cat sat on the mat." The root node is *Sentence*, which branches into *NP* (The cat) and *VP* (sat on the mat). The *NP* branches into *The* and *cat*. The *VP* branches into *sat*, *PP* (on the mat), and *NP* (the mat). The *PP* branches into *on* and *NP* (the mat). The *NP* (the mat) branches into *the*, *mat*, and *NP* (the mat). The *NP* (the mat) branches into *the*, *mat*, and *NP* (the mat).





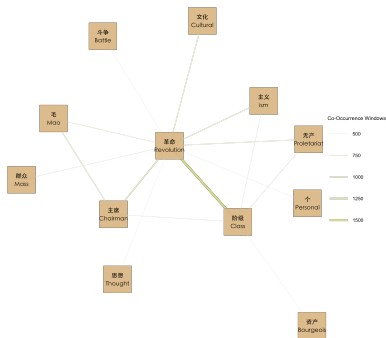
# The Keyword Extraction Scaling Approach

The Tokens of Part-of-Speech Tags Reduced to Noun, Verb and Adjective



**Figure:** An examples of grammar annotation reduced to noun, adjective and verbon Universal Dependencies. It also includes major syntactic relations such as clausal complement (ccomp), adverbial clause modifier (advcl), nominal subject(nsubj). Note: This figure presents an example of semantic role annotations based on Universal Dependency implementation using single sentence from the Big-character poster. The original sentence is 只要她反对毛主席，搞自己的一套，阳奉阴违，搞神秘化，我们就要反对她，揪出她!.

# The Keyword Extraction Scaling Approach



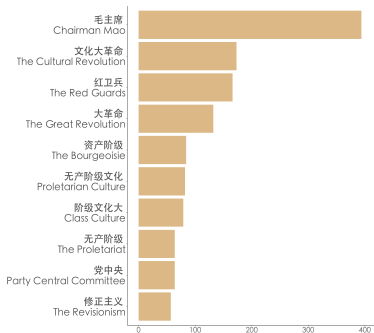
(a) The First Marxist-Leninist Wallposter



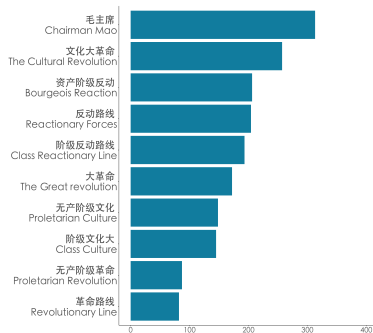
(b) The Zhou Enlai's Declaration

**Figure:** Visualisation of the co-occurrences for the top 15 most frequent co-occurring nouns and adjectives from documents within major historical incidents (translations given in text).

# The Keyword Extraction Scaling Approach



(a) The First Marxist-Leninist Wallposter



(b) The Zhou Enlai's Declaration

**Figure:** Four examples of the top 10 keyword phrases identified by TextRank from documents within major historical incidents (translations given in text).

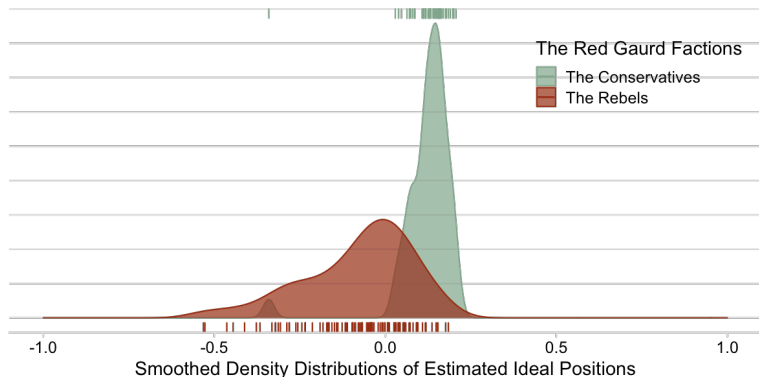
# The Keyword Extraction Scaling Approach

## Slogan-based Wordfish (Sworfish) Scaling Model:

- The algorithm estimates positions of each faction based on **the Poisson scaling model** proposed by Slapin and Proksch (2008).
- The model produces the estimated positions by **analyzing frequencies of extracted keywords and phrases** from dictionaries constructed in the first stage.
- The model takes the following functional form:

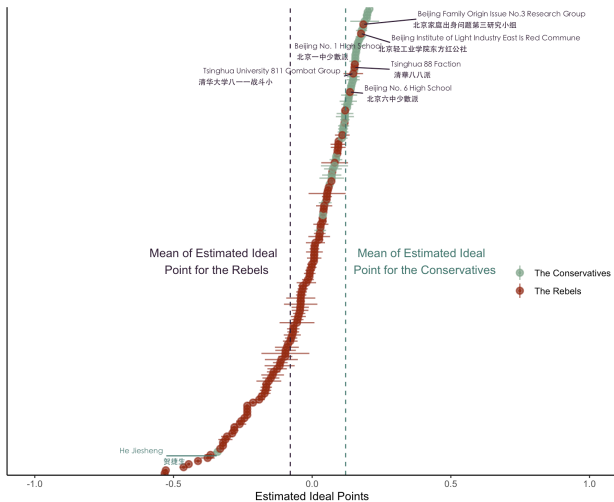
$$\begin{aligned}y_{jm} &\sim \text{Poisson}(\theta_{jm}) \\ \theta_{jm} &= \exp(a_m + b_j + \beta_j * w_i^m)\end{aligned}\tag{1}$$

# The Distribution of Factions



**Figure:** Smoothed density distributions of estimated positions for the Red Guard participants

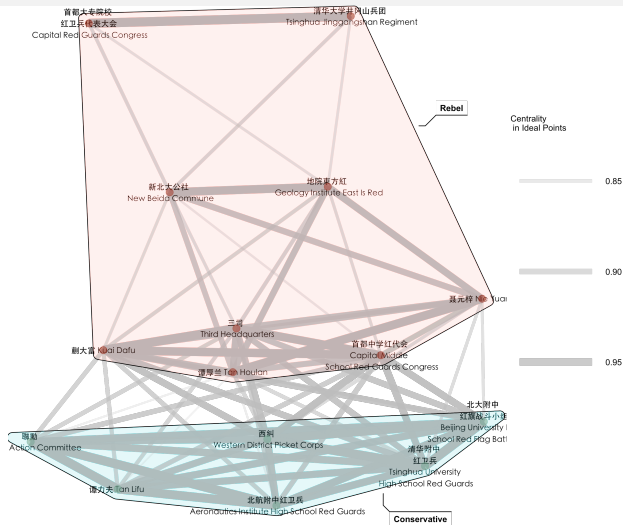
# Individual Ideal Points



David, Yen-Chieh Liao, Yi-Nung Tsai, Daniel Tene, and Dechun Zhang

Factionalism and the Red Guards under Mao's China

# Foremost Representative Factions



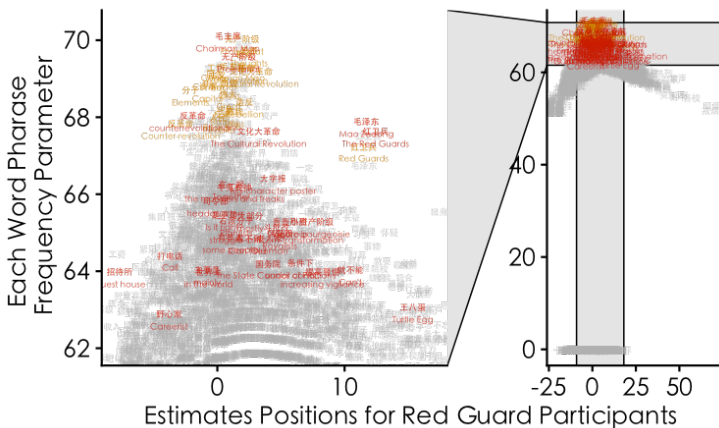
David, Yen-Chieh Liao, Yi-Nung Tsai, Daniel Tene, and Dechun Zhang

Factionalism and the Red Guards under Mao's China

MIT



# Most Frequent Slogans Discrimination Parameters



David, Yen-Chieh Liao, Yi-Nung Tsai, Daniel Tene, and Dechun Zhang

Factionalism and the Red Guards under Mao's China

# Most Frequent Slogans Discrimination Parameters

Table: Top 10 Keyphrases (and Keywords) Placing Participants on Single Dimension

<b>Fish Head</b> Less Weights	<b>Pelvic (Upper)</b> Large Weights on Positive	<b>Dorsal (Lower) Fin</b> Large Weights on Negative
毛主席 (Chairman Mao), 文化大革命 (The Cultural Revolution), 无产阶级 (Proletariat), 反革命 (Counter-Revolution), 牛鬼蛇神 (Dark Ghosts), 右派分子 (Rightists), 罪魁祸首 (The Culprit), 打下去 (Fight Down), 大学生 (College Students), 司令部 (Headquarters), 人民政府 (People's Government), 无政府主义 (Anarchism), 中学生 (Middle School Student), 抓起来 (Catch Up), 革命党人 (Revolutionaries), 反革命份子 (Counter-Revolutionaries), 极右派 (Ultra Rightists)	红卫兵 (Red Guards), 毛泽东 (Mao Zedong), 小资产阶级 (petty bourgeoisie), 唯物辩证 (Dialectical Materialism), 民主集中制 (Democratic Centralism), 辩证唯物 (Dialectical Materialism), 革命烈士 (Revolutionary Martyrs), 臭气冲天 (Smelly Stinking), 解放思想 (Emancipating the Mind), 昏迷不醒 (Unconscious), 特权阶层 (Privileged Class), 打天下 (struggle to seize state power), 接班人 (Successor), 小学生 (Primary School Student),	极左派 (Ultra Leftist), 炊事员 (cooker), 胰岛素 (Insulin), 女演员 (Actress), 极左派 (Ultra Leftist), 专程去 (Make a special trip), 一起去 (Go Together), 贺龙 (He Long), 病人 (Patient), 治疗 (therapy), 招待 (serve), 享受 (enjoy), 诊断 (diagnosis), 奢侈 (luxury), 叛国 (treason), 津贴 (allowance), 高薪 (high salaries), 专家 (Specialist), 信忠 (Xinzong), 提审 (arraignment), 待遇 (treatment), 案件 (case)

## Takeways

- This paper contributes a new strain of text scaling method (Slogan-based Wordfish, Swordfish) that strategically combines the TextRank algorithm and “Wordfish”.
- We construct a bag of keyword-based collocation based on the Red Guard’s publications. Our estimation strategy unveils better semantic themes in a given background.
- This is particularly true for Mandarin, a language with no spaces, where the same word can frequently have different meanings depending on the context in which it resides.
- Our finding echoes evidence from the number of early qualitative research on the Cultural Revolution.

Thank you !!

Our replication codes and full paper can be found at the authors' Github, <https://github.com/davidycliao/redguards>.