

Roberto Franzosi

Emory University

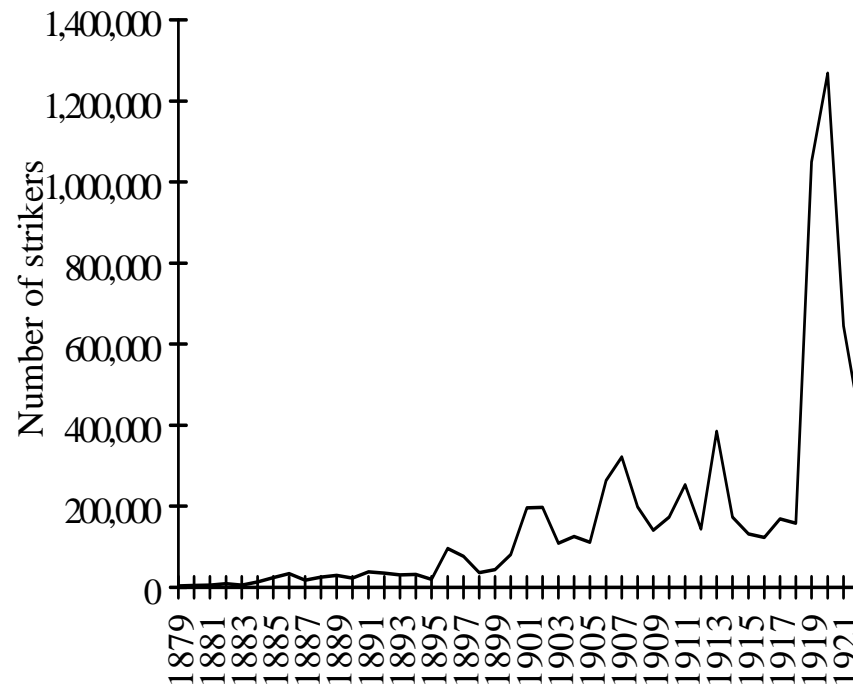
Relations in Texts
(and How to Get at Them)

Conference on Optimal Coding of Open-Ended Survey Data
ISR, University of Michigan
December 4-6, 2008

Which relations in texts?

- Between words
- Between concepts
- Between social actors and their actions

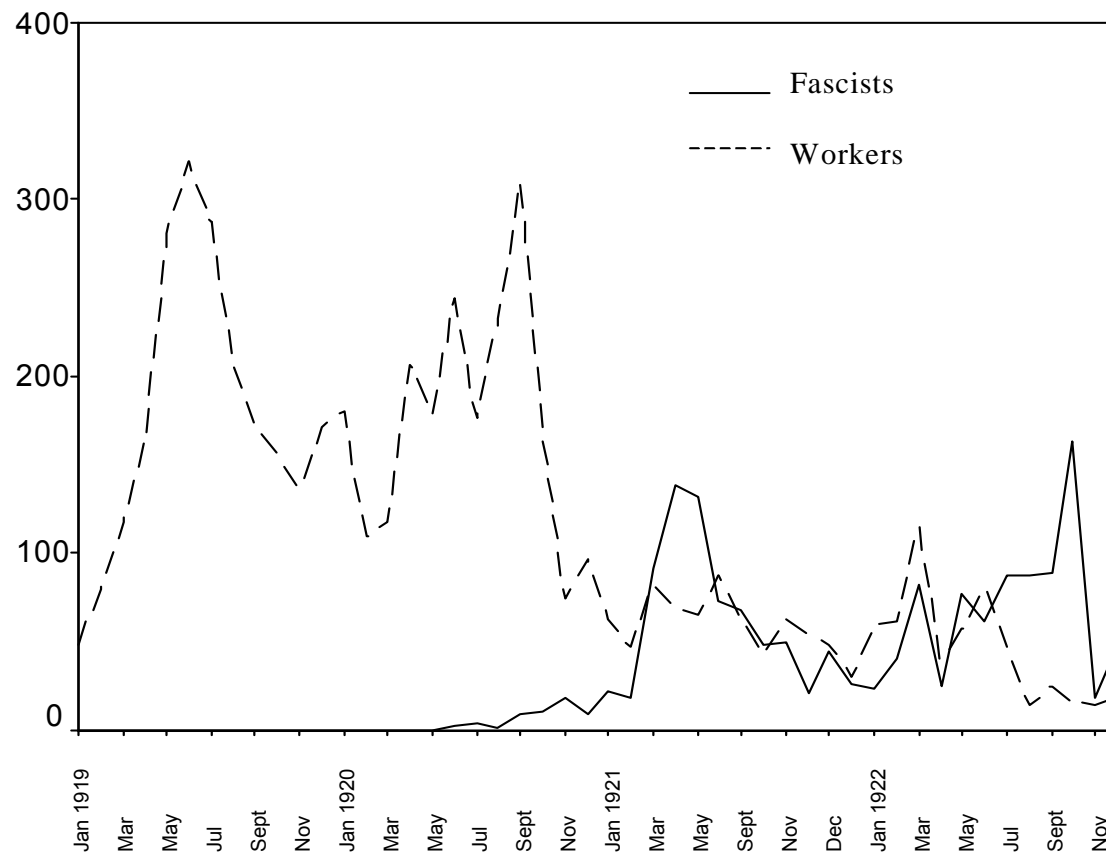
Why the 1919-22 period?



Monthly number of strikers (1879-22)

“Red Years” and “Black Years” (1919-22)

Working-class mobilization and fascist counter-mobilization

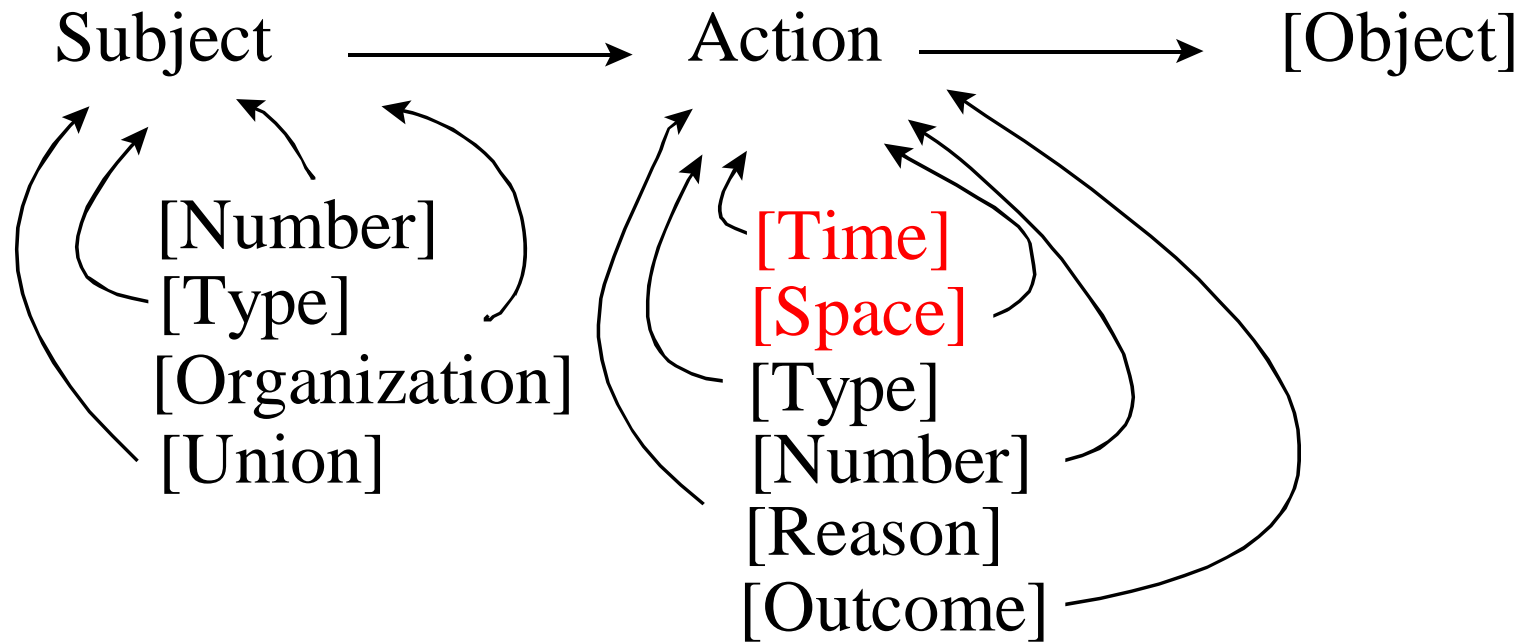


Typical text

Fascists arrive on a truck on the night of 5/7/1921 in Bissone di S. Cristina at the pub of Mr. Prati. There, they seize 5 workers and take them outside. They beat up the 5 workers with retorted cowhide with lead inside.

Story grammar

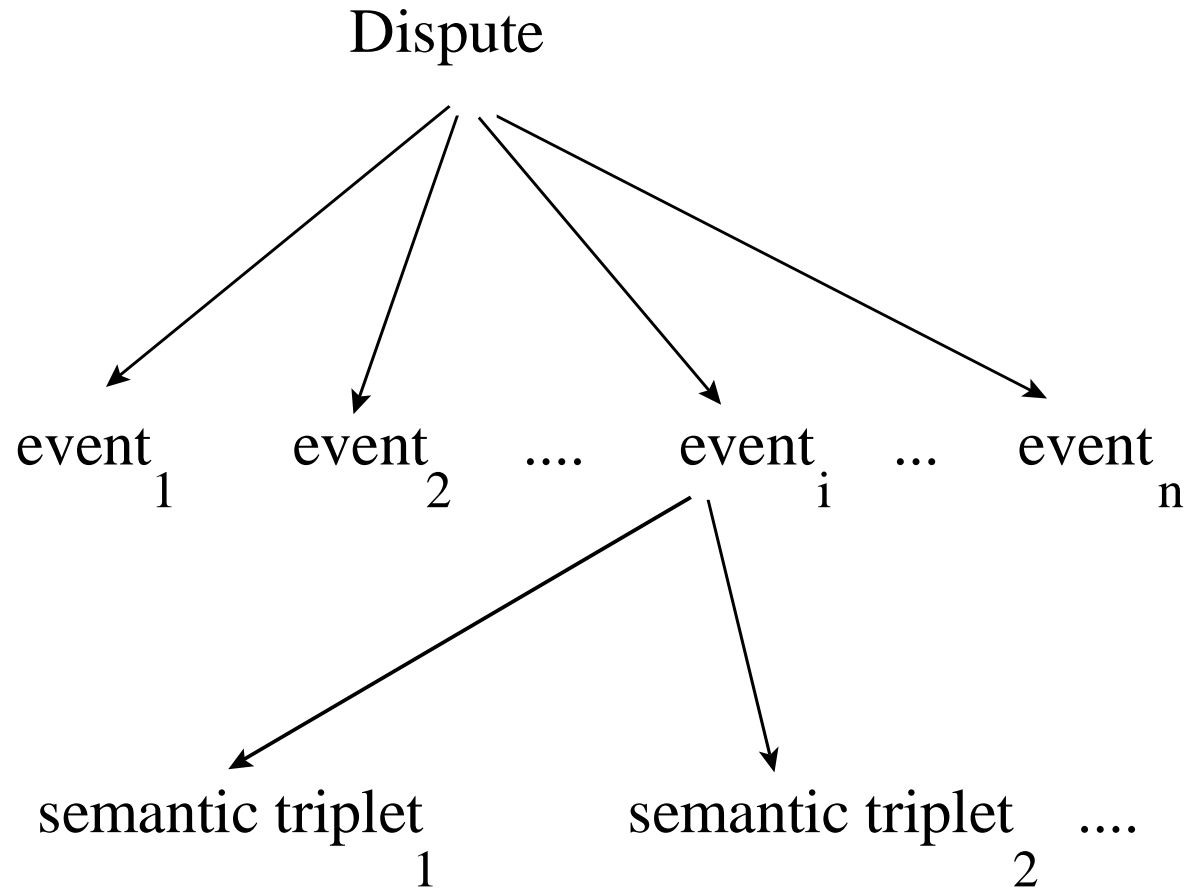
(“semantic triplet”)



Rewrite rules

<triplet> → {<subject>} {<action>} [{<object>}]
<subject> → {<actor>} [{<modifier>}]
<actor> → {<individual>} | {<set of individuals>} |
{<institution>}
...
<modifier> → [{<type>} [<organization>] [<number>]
...
<type> → (male | young | immigrant | skilled | ...)
...

Hierarchical structure



PC-ACE

Program For Computer-Assisted Coding of Events
(available for free download at www.pc-ace.com)



My data: PC-ACE coded output

Semantic triplet: [Subject: [Actor... (Name of collective actor: **FASCISTS**)]]
[Action: (Verb phrase: **ARRIVE**) [Time... (Definite date: **05/07/1921**)]
[Space..City: (City name: **BISSONE DI S. CRISTINA**) (Locality: [Physical
object: (Name of physical object: **PUB**) [Ownership...Individual: (Name of
individual actor: **INDIVIDUAL**) [Personal characteristics... (Last name:
PRATI)]]]]

...

Semantic triplet: [Subject: [Actor... (Name of collective actor: **FASCISTS**)]]
[Action: (Verb phrase: **SEIZE**) [Object: [Actor..Collective actor: (Name of
collective actor: **WORKERS**) [Collettive characteristics.... (Single value: **5**)]]]]

Semantic triplet: [Subject: [Actor... (Name of collective actor: **FASCISTS**)]]
[Action: (Verb phrase: **TAKE OUTSIDE**) [Action: (Verb phrase: **BEAT UP**)
[Instrument.. (Type of instrument: **RETORTED COWHIDE WITH LEAD**)]
[Object: [Actor..Collective actor: (Name of collective actor: **WORKERS**)
[Collettive characteristics.... (Single value: **5**)]]]]

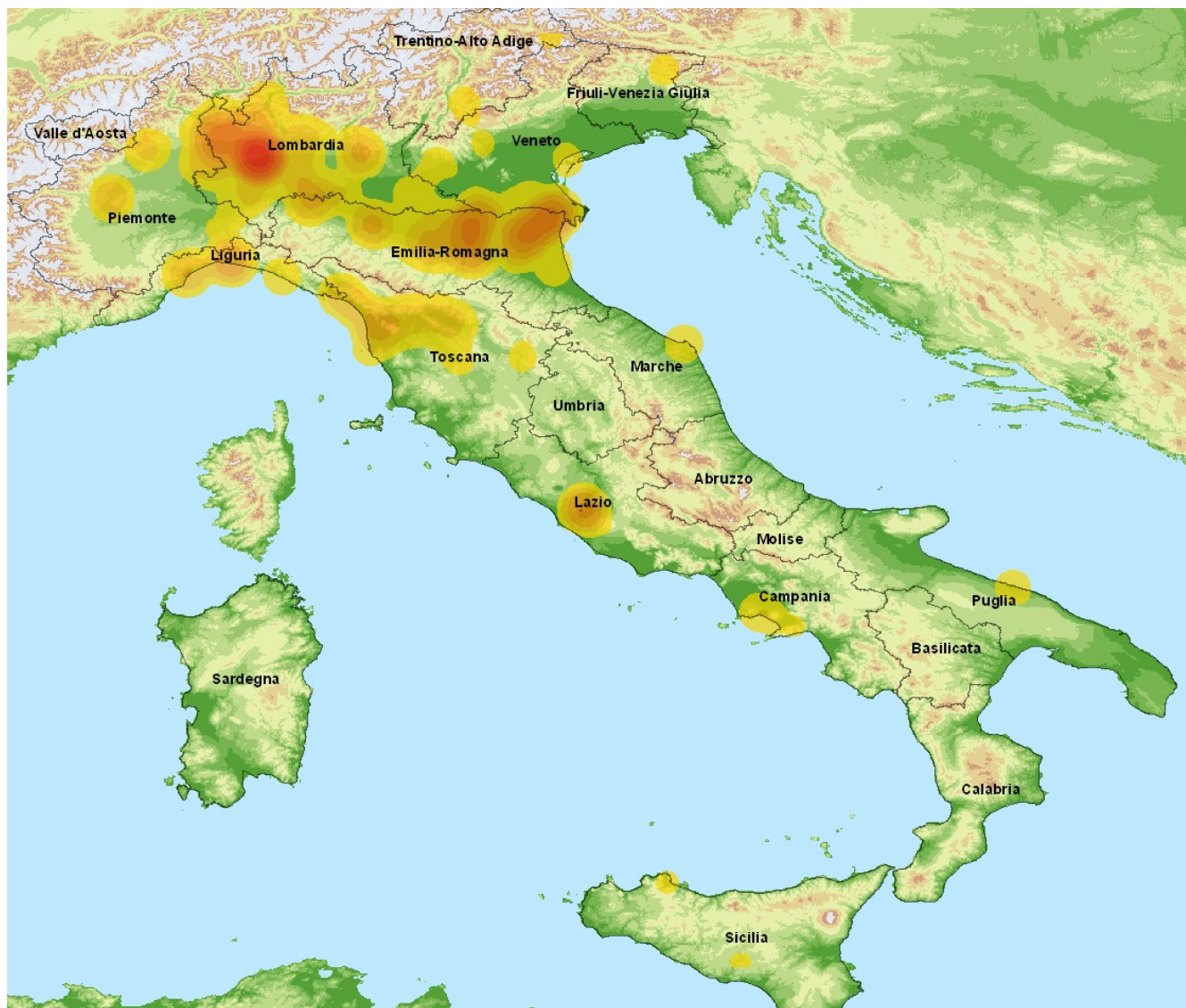
Sources and data

Newspaper name	<i>Il Lavoro</i>	<i>Avanti!</i>	<i>Popolo d'Italia</i>
No. of articles	15,400	17,894	19,871
No. of triplets	17,601	139,530	93,558
Year completion of data collection	1990	2006	2008

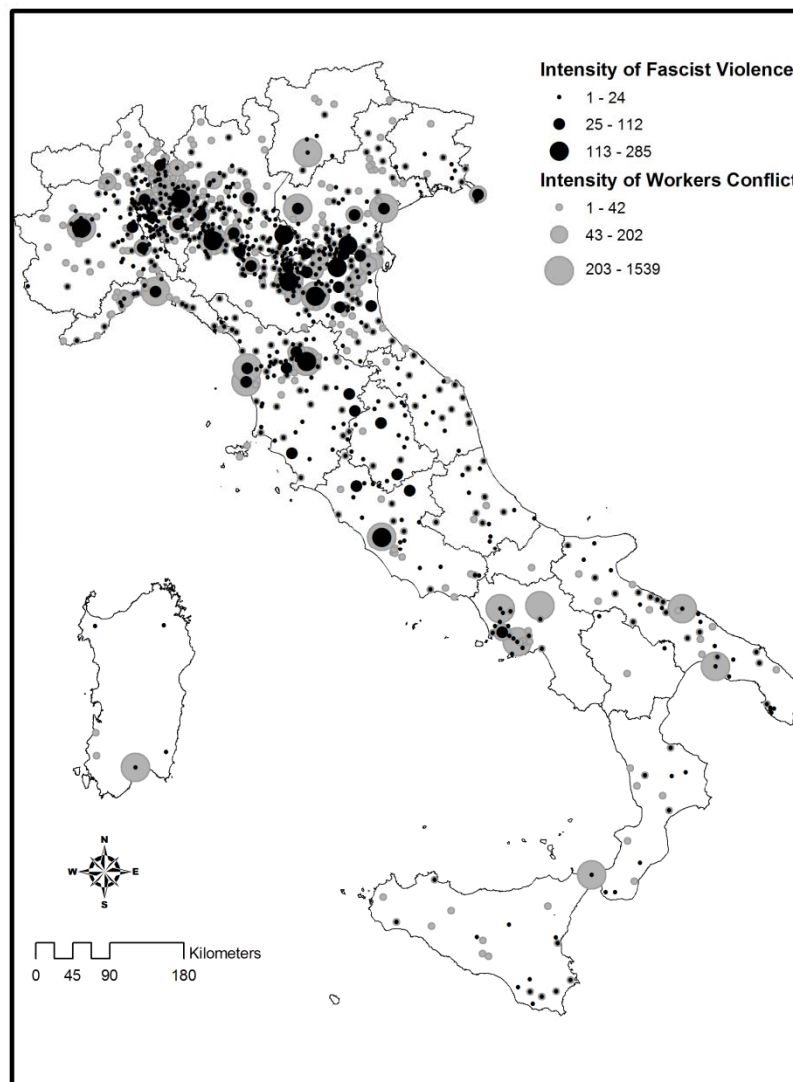
Acknowledgments

- 1.University of Wisconsin-Madison
- 2.National Science Foundation
- 3.University of Reading
- 4.University of Trento
- 5.Provincia Autonoma di Trento
- 6.Emory University

Geography of *Conflict & Violence*

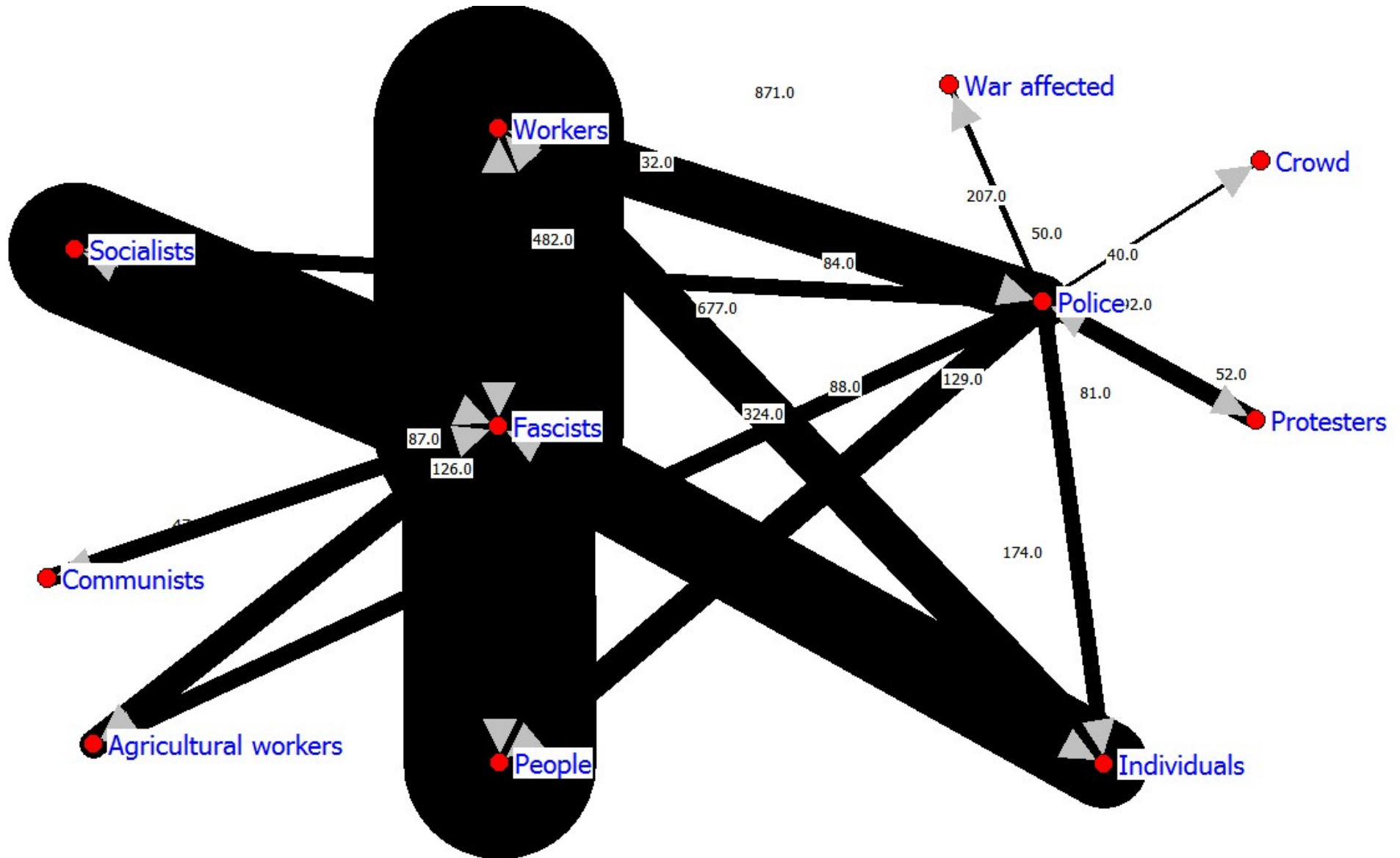


Geography of *Conflict & Violence*



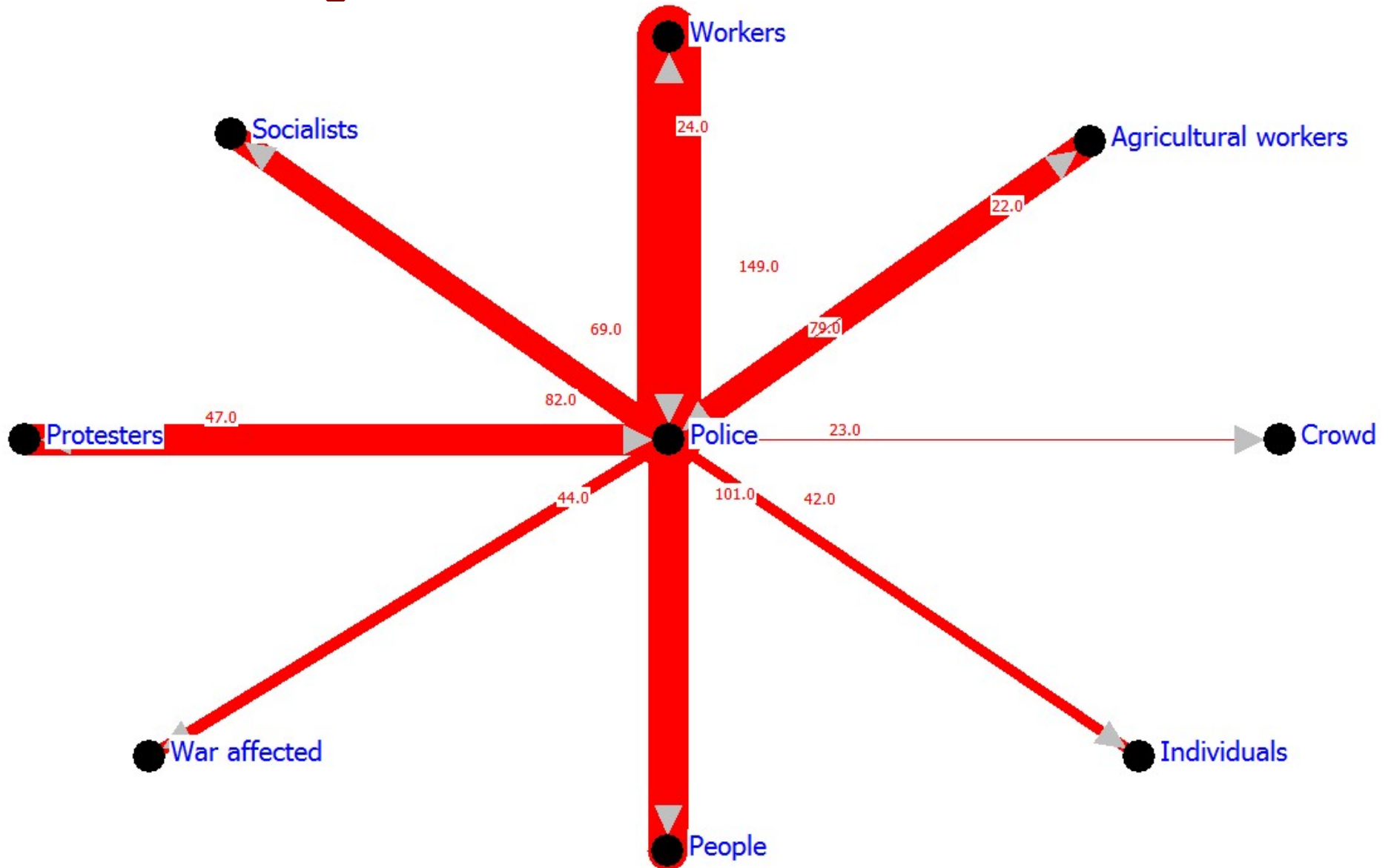
1919-1922: A bird's eye view

Two centers of violence: Police & Fascists



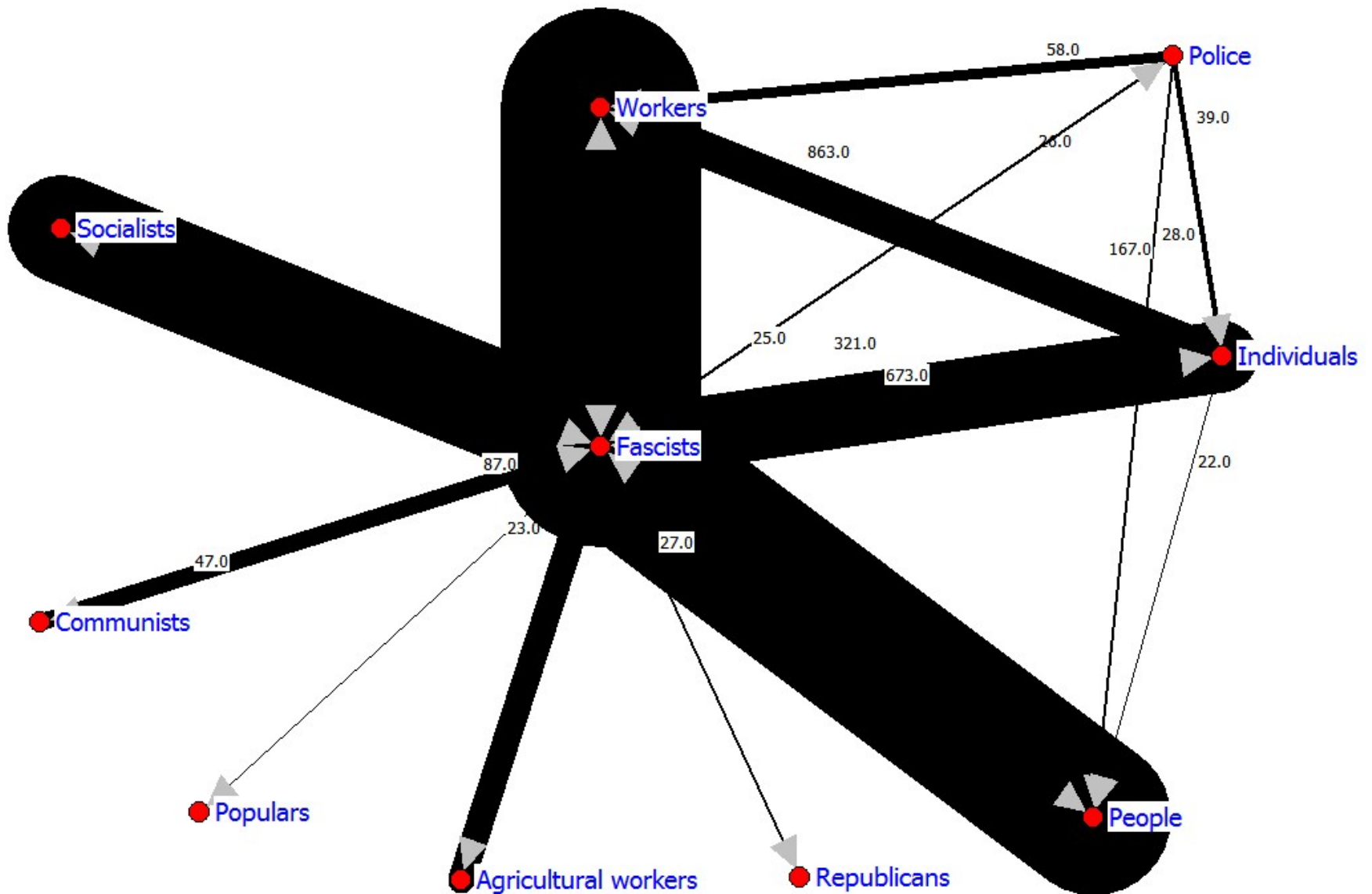
The “red years” (1919-1920)

Working-class mobilization: Police & violence

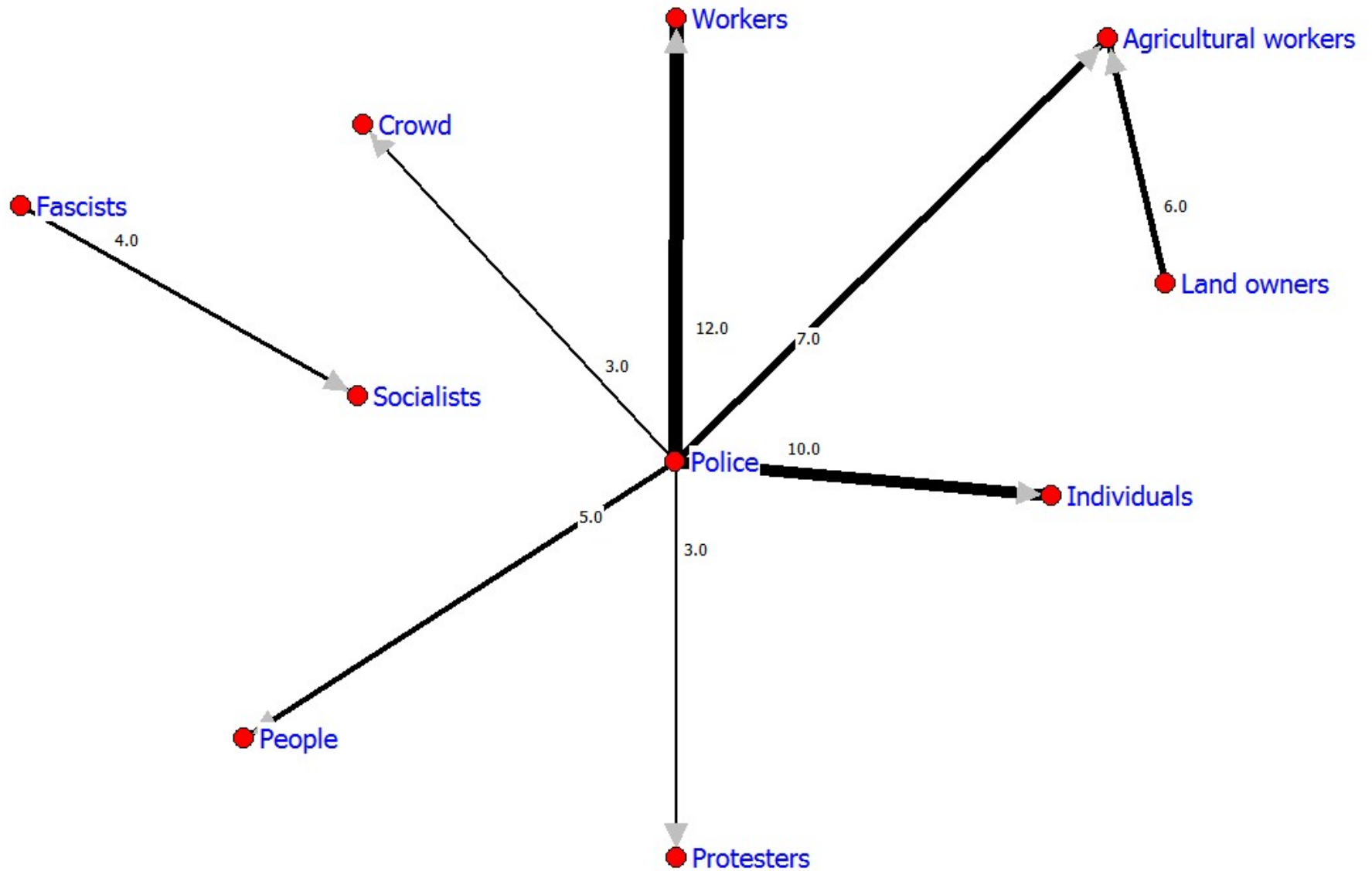


The “black years” (1921-1922)

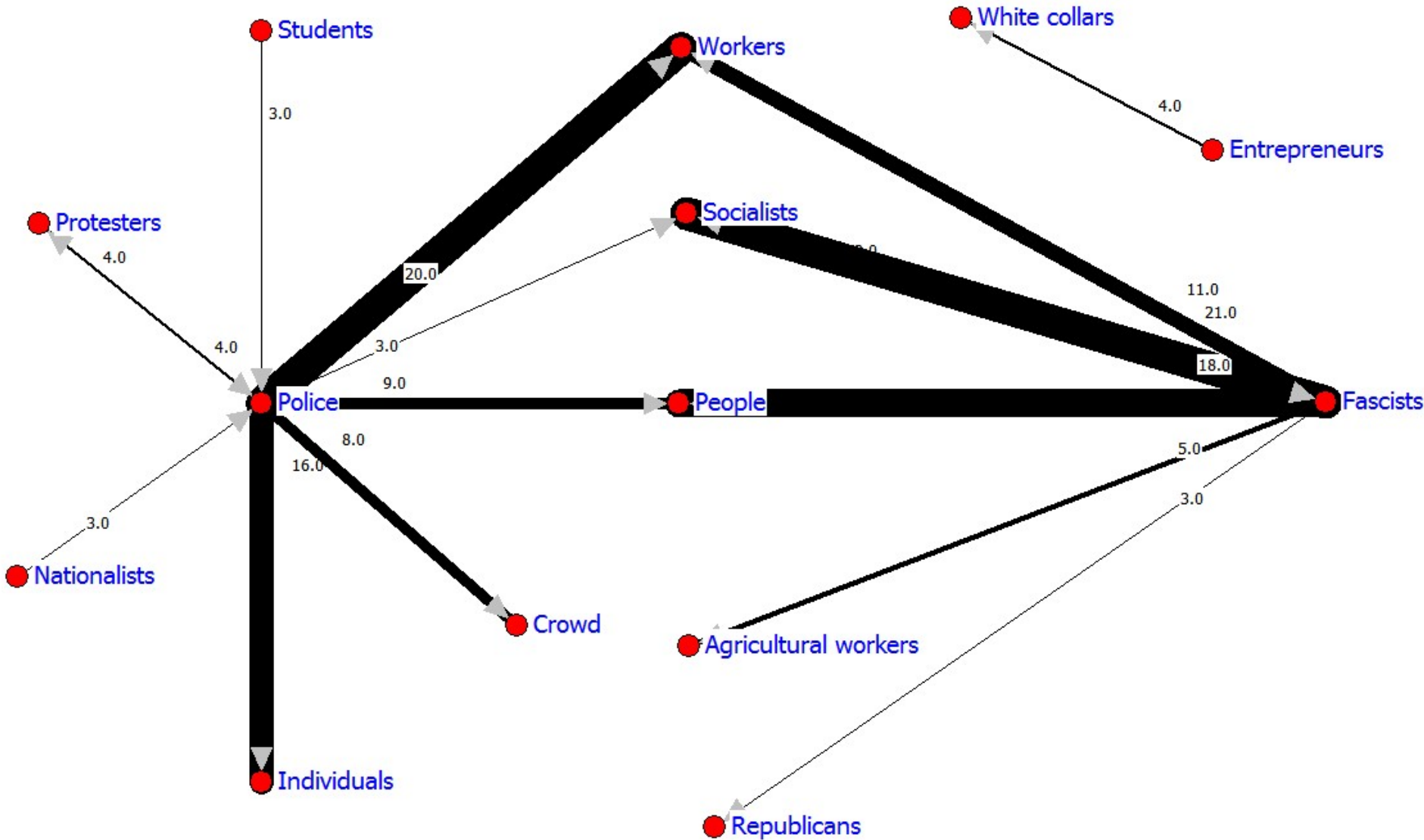
Counter mobilization: Fascists & violence



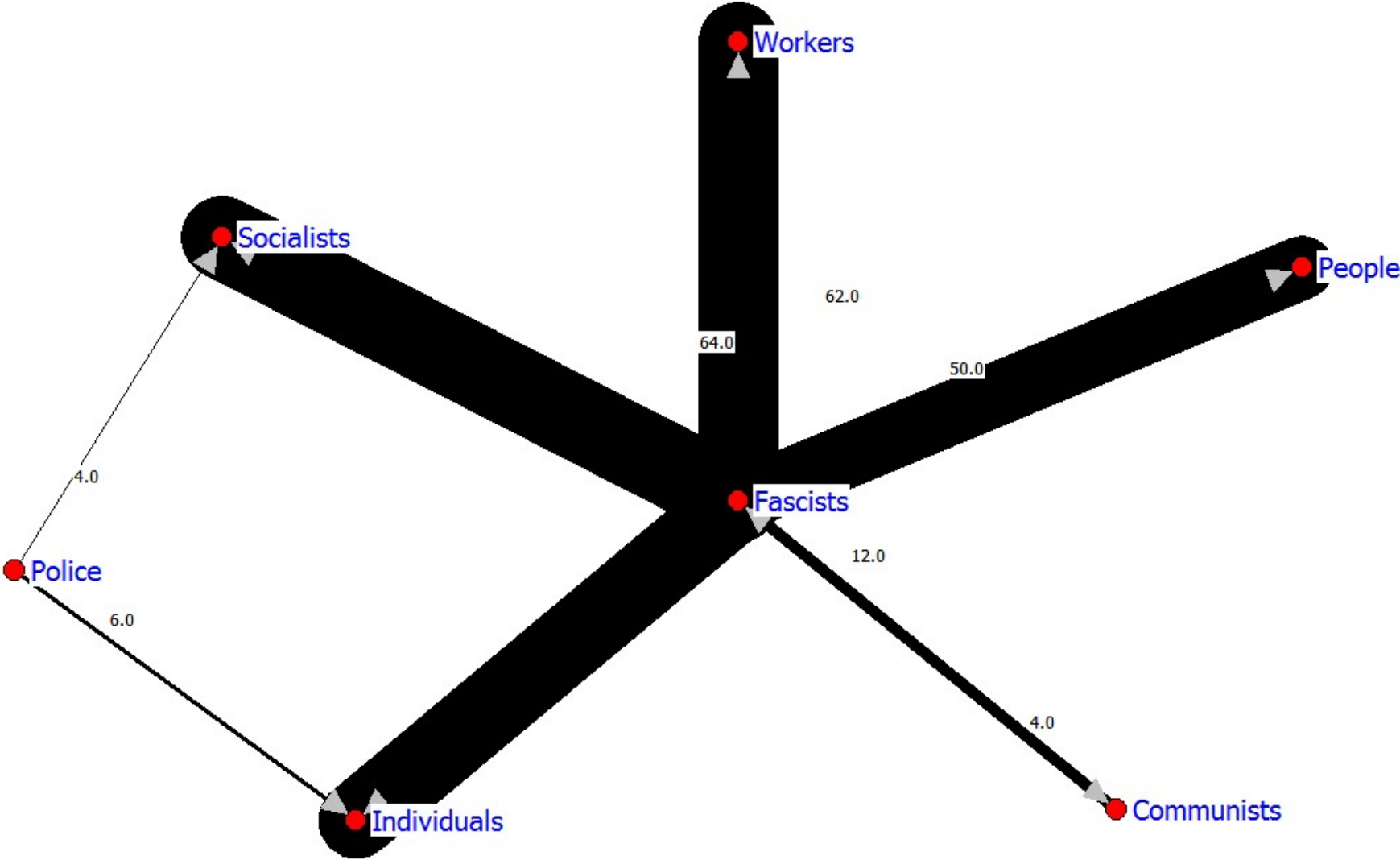
1920, 3rd Quarter (Factory Occupation)



1920, 4th Quarter



1921, May



Pros & Cons

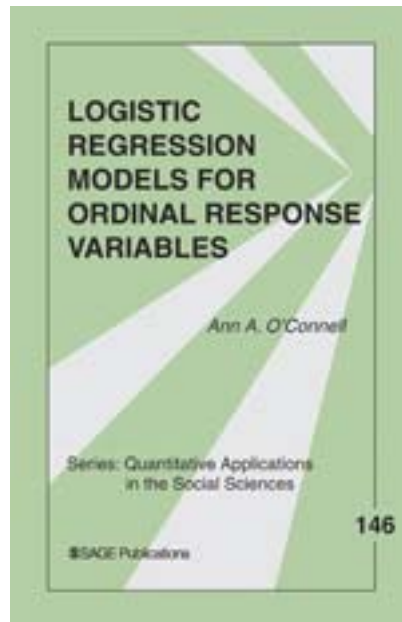
- **Pros**

- Linguistic theory in coding scheme design
- Nearly hypothesis-free data
- Higher data reliability (semantic coherence)

- **Cons**

- Works well with narrative type of texts only
- Very labour intensive

Quantitative Narrative Analysis (Sage, 2009)



SAGE QASS SERIES

Quantitative Applications in the Social Sciences

The end