Polarization Games over Social Networks

Xilin Zhang, Emrah Akyol, Zeynep Ertem

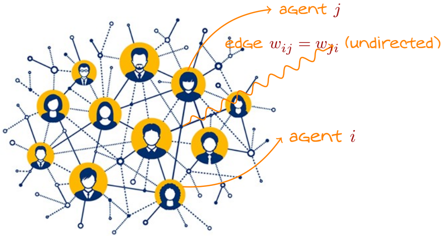
Binghamton University
This work is supported in part by NSF grant CCF (CAREER) # 2048042

June 10, 2024

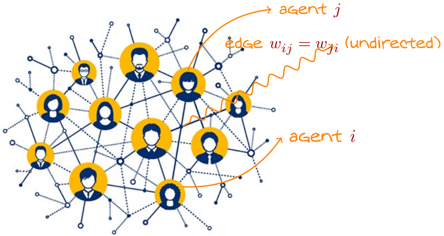
What is a Polarization Game?

- A game with two players with opposing objectives: a minimizer and a maximizer of polarization
- Polarization $\Rightarrow$ sample variance of expressed opinions of all individuals in a network
- Individuals are in a network, their expressed opinions evolve in time via known dynamics (this work: Johnson-Friedkin model)
- What can players change? Innate opinion of a selected agent. Action set: each player chooses one agent and change its innate opinion.
- Inherently a zero-sum game

Problem Setting



Problem Setting



Opinion dynamics: Friedkin-Johnson

expressed opinion

innate opinion

$$z_i(t) = \alpha_i s_i + \sum_{j \in V} w_{ij} z_j(t-1)$$

$\alpha_i \in [0, 1]$

network influence

Problem Setting

Opinion dynamics: Friedkin-Johnson

expressed opinion

innate opinion

$$z_i(t) = \alpha_i s_i + \sum_{j \in V} w_{ij} z_j(t-1)$$

$\alpha_i \in [0, 1]$

network influence

$\mathbf{s} = (s_i, s_j, s_{-(i,j)})$
original

Maximizer$[i^A, s_{i^A}]$ →$[i, s'_i]$ | Minimizer$[i^B, s_{i^B}]$ →$[j, s'_j]$

$\mathbf{s}' = (s'_i, s'_j, s_{-(i,j)})$
modified

Opinion dynamics →

$\mathbf{z}' = \lim_{t \to \infty} \mathbf{z(t)}$
steady-state opinion

→

$P(\mathbf{z}') = (\mathbf{z}' - \bar{\mathbf{z}})^T (\mathbf{z}' - \bar{\mathbf{z}})$
polarization

Nash vs. Generalized Nash

- What happens if both players select the same individual?
- Two options, yielding to two different game formulations
  - Both players get zero payoffs (zero-sum: This talk's focus)
  - We externally enforce that this does not happen, i.e., action sets are constrained to be disjoint (hard to track, but more realistic: Generalized Nash, non-zero sum– each player gets $-\infty$ payoff $\rightarrow$, general sum, our other work).
- Network information, opinion evolution dynamics, objectives are common knowledge (available at both players).

Nash Equilibrium

A tuple (i_A, s_{i^A}) and (i_B, s_{i^B}) is a NE if:

$$i_A, s_{i^A} = \underset{i_A \in \mathcal{S}_A, s_{i^A} \in [0,1]}{\arg\max} P,$$

$$i_B, s_{i^B} = \underset{i_B \in \mathcal{S}_B, s_{i^B} \in [0,1]}{\arg\min} P$$

where $P = P_0$ (pre-game polarization) if $i_A = i_B$.

Literature Review

- Two relevant works:
- Chen-Razc. An adversarial model of network disruption: Maximizing disagreement and polarization in social networks. IEEE Trans. Network Science, 2021 (maximizing polarization via innate opinion change, same network, polarization, and opinion dynamics model)
- Zhu et.al. Minimizing polarization and disagreement in social networks via link recommendation. NeuroIPS 2021 (minimizing polarization via edge-removal, same polarization and opinion dynamics model)
- No prior work in the game framework where a minimizer and a maximizer exists.

Algorithms to find NE

- We use fictitious play (FP) to obtain NE
- Fictitious play: Find the best pure response (deterministic: one individual and one innate opinion) that maximize/minimize polarization assuming other player is playing a mixed strategy (stochastic: probabilities over individuals, innate opinions) based on its history. FP converges to NE for zero-sum games.
- We use following with exhaustive search i_A, i_B over the entire network.

Theorem

For a given i_A, i_B, pair, $s_{i^A} \in \{0, 1\}$ and

$$s_{i^B} = \frac{-\sum_{j \neq i_B} s_j (a_j - \frac{1}{n})^T (a_i - \frac{1}{n})}{(a_i - \frac{1}{n})^T (a_i - \frac{1}{n})}$$

where a_i is the i'th column of $A = (I + L)^{-1}$ and L is the Laplacian of the network.

- Proof: Note $\mathbf{z}' = (I + L)^{-1}\mathbf{s}$ and $P(\mathbf{z}')$ is convex in $\mathbf{z}'$. SP is linear in $\mathbf{s}$, we are maximizing/minimizing a convex functional. Maximizer is on the boundary and the minimizer

Numerical Results

- The rest of the presentation is empirical, to get an intuition over how NE behaves.
- Algorithms are generally too computationally complex.

Numerical Results

- The rest of the presentation is empirical, to get an intuition over how NE behaves.
- Algorithms are generally too computationally complex. Can we get away with some heuristics on choosing the individual based on empirical analysis?
- Our observations lead to the following heuristics:
  - maximizer $\Rightarrow$ Location, location, location! Maximizer chooses the least connected (lonely!) agent, regardless of the innate opinion distribution of the network to avoid getting mitigated via the averaging process.

Numerical Results

- The rest of the presentation is empirical, to get an intuition over how NE behaves.
- Algorithms are generally too computationally complex. Can we get away with some heuristics on choosing the individual based on empirical analysis?
- Our observations lead to the following heuristics:
  - maximizer $\Rightarrow$ Location, location, location! Maximizer chooses the least connected (lonely!) agent, regardless of the innate opinion distribution of the network to avoid getting mitigated via the averaging process.
  - minimizer $\Rightarrow$ extreme opinion individual while also paying attention to the network location.
- We next analyze empirically how good this heuristic rule of thumb is.

Observations

- Polarization tends to increase in both networks due to the polarization game, as predicted by theory.
- The steady-state opinions tend to cluster narrowly around the mean due to averaging dynamics in opinion formation.
- The Minimizer has limited scope to reduce polarization due to the already low variance around the mean.
- The Maximizer can significantly increase polarization by pushing selected agents' opinions to extreme values (0 or 1).
- Less connected agents are typically targeted by the Maximizer to maximize polarization impact by reducing the averaging effect.
- The Minimizer often selects the same agents as the Maximizer to minimize the Maximizer's impact on polarization.
- In networks like Reddit, where opinions are more centered around the mean compared to networks like KC, the Maximizer finds more opportunities to influence and increase polarization, as indicated by a larger observed η value.

Three modes of network dynamics

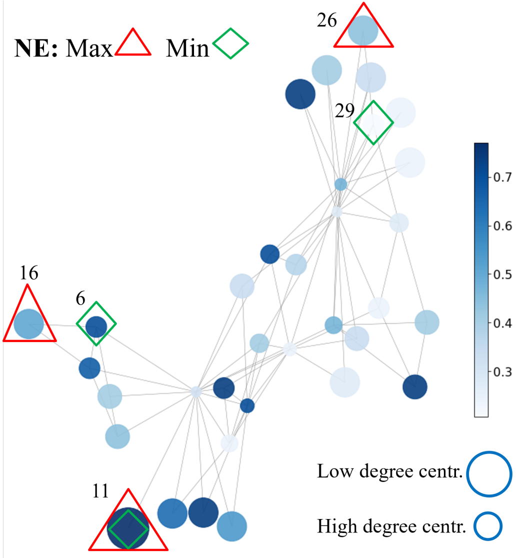
ICC Paper network dynamics

- Opinion dynamics
  - Agents in the network exchange opinion according to F-J model

Journal Extension

- No opinion dynamics
  - Agents in the network do not exchange opinion
- Stubborn agent in F-J opinion dynamics
  - Players' chosen agents will be the stubborn agents in the network

Opinion Dynamics - KC Network

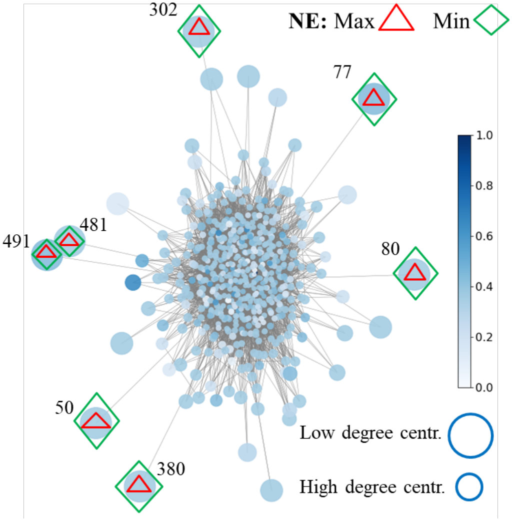


Data	Size	Network	Polariz.	Source
Karate club	34	undirected	0.16	(1)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	11	0.77	1	0.6	11	0.77	0.39	0.11	1/8
	26	0.45	0	0.13	6	0.72	0	0.86	
	16	0.51	1	0.27	29	0.21	1	0.03	

Table 1 Games on KC Network.

Opinion Dynamics - Reddit Network



Data	Size	Network	Polariz.	Source
Reddit	553	undirected	0.01	(2)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	50	0.5	1	0.14	50	0.5	0.5	0.14	11
	77	0.5	1	0.14	77	0.5	0.5	0.14	
	80	0.56	1	0.15	80	0.56	0.5	0.15	
	302	0.5	0	0.14	302	0.5	0.5	0.14	
	380	0.5	1	0.14	380	0.5	0.5	0.14	
	481	0.5	1	0.15	481	0.5	0.5	0.15	
	491	0.56	1	0.14	491	0.56	0.5	0.14	

Table 2 Games on Reddit Network

Opinion Dynamics - Opinion Distribution

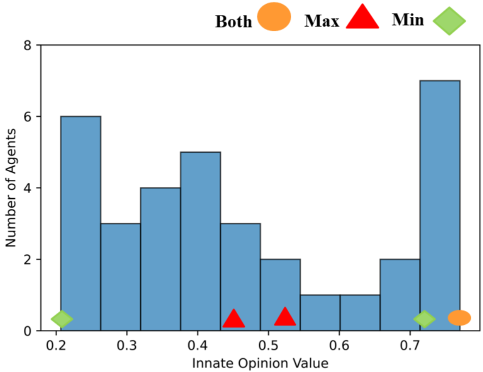


Figure 1 KC opinion distribution

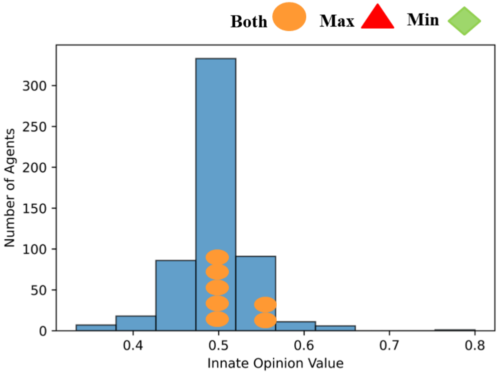
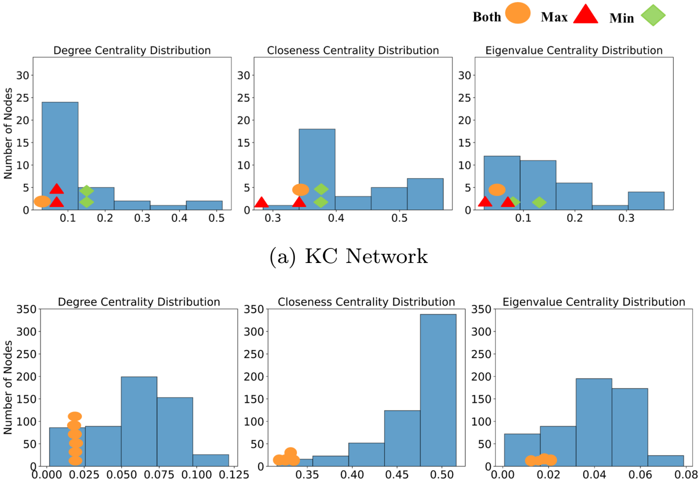


Figure 2 Reddit opinion distribution

- KC - Minimizer prefers nodes with extreme opinion, Maximizer prefers nodes with a neutral opinion
- Reddit - The network is large, and the effects can be averaged out if choosing any nodes with a lot of neighbors regardless of its innate opinion.

Opinion Dynamics - NE Centrality Property



(a) KC Network

(b) Reddit Network

Figure 3 Nodes centrality distribution in network

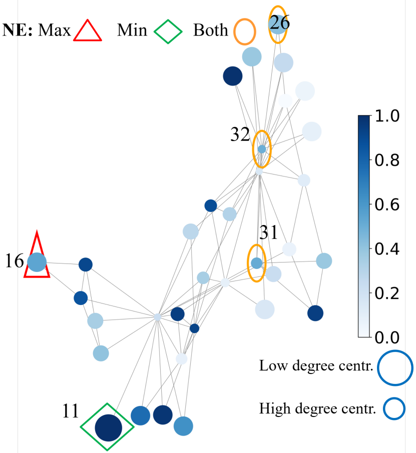
- KC - Max prefers lonely agents with less centrality, F-J is an averaging process, that can cancel out Max's effects
- Reddit - Both Max and Min prefer lonely agents, Min is going after Max to cancel out its effects

Conclusion

- We have empirically studied polarization games, to the best our knowledge, this is the first work to study polarization games.
- Analyzed the functional properties of zero-sum games including the optimal responses for players and properties of Nash equilibria.
- Real network simulations match theoretical results on player behaviors in zero-sum games.
- Empirical modeling of polarization dynamics over networks and the analysis of equilibrium strategies for external players aiming to maximize or minimize polarization.
- Extension: Various network dynamic modes, such as static network, set information sources in the network

Extension

No Opinion Dynamics - KC



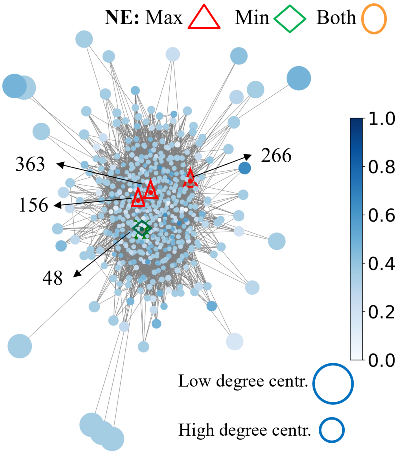
Data	Size	Network	Polariz.	Source
Karate club	34	undirected	0.16	(1)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	16	0.51	1	0.07	11	0.77	0.58	0.99	
	26	0.45	1	0.31	26	0.45	0.73	0.003	
	31	0.51	1	0.31	31	0.51	0.79	0.003	
	32	0.51	1	0.31	32	0.51	0.80	0.004	

Table 3 No Op Dynamic on KC Network.

- Location doesn't matter
- Max: All chosen agents have neutral innate op.
- Choose 26, 32, 33 with almost equal probability
- Min: 99% probability choosing most extreme opinion(node 11) in the network
- Other actions to counteract Max's effect (negligible probability)

No Opinion Dynamics - Reddit



Data	Size	Network	Polariz.	Source
Reddit	553	undirected	0.01	(2)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	156	0.5	1	0.36	48	0.8	0.5	1	
	266	0.5	1	0.36					
	363	0.5	1	0.28					

Table 4 No Op Dynamic on Reddit Network.

- Location doesn't matter
- Max chooses agents with a neutral opinion
- Min chooses 48 - the biggest opinion in the network

No Opinion Dynamics - NE Opinion Property

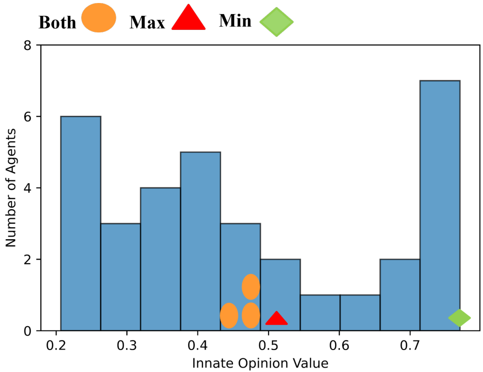


Figure 4 KC opinion distribution

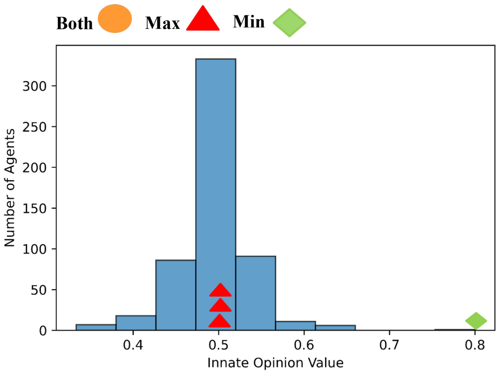
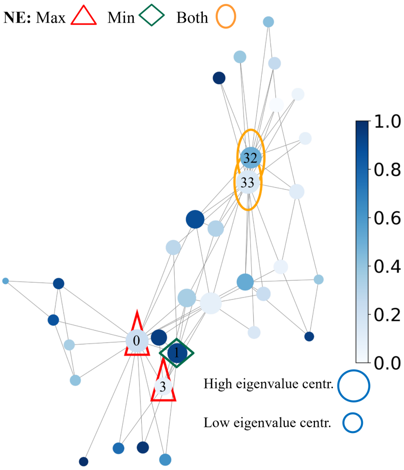


Figure 5 Reddit opinion distribution

- KC - Min prefers nodes with extreme opinion, Max prefers nodes with a neutral opinion
- Reddit - No opinion dynamic, players' actions can be effective regardless of the node's location. Max chooses neutral ops, Min selects the only outlier opinion

Stubborn Agent - KC



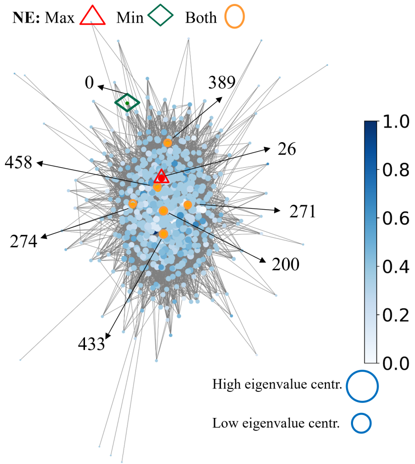
Data	Size	Network	Polariz.	Source
Karate club	34	undirected	0.16	(1)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	32	0.49	1	0.3	32	0.49	0.72	0.006	
	33	0.29	1	0.20	33	0.29	0.44	0.36	
	0	0.33	1	0.14	1	0.73	0.12	0.24	
	3	0.26	1	0.37	1	0.73	0.13	0.4	

Table 5 Stubborn game on KC network.

- All chosen nodes are high influential in the network
- Min: 1 - high influential node with extreme opinion, the neighbor with node 0 and 3
- Max: 0 and 3 - high influential nodes with a neutral opinion
- Min choose node 32 with negligible prob., its innate op is neutral

Stubborn Agent - Reddit



Data	Size	Network	Polariz.	Source
Reddit	553	undirected	0.01	(2)

	Maximizer				Minimizer				
	i^1	s_i	s_i^1	p	i^2	s_i	s_i^2	p	η
NE	200	0.49	1	0.008	200	0.49	0.03	0.0005	
	271	0.44	1	0.95	271	0.44	1	0.001	
	274	0.49	1	0.008	274	0.49	0.04	0.0001	
	389	0.49	1	0.008	389	0.49	0.04	0.0003	
	433	0.42	1	0.009	433	0.42	0.03	0.0003	
	458	0.50	1	0.008	458	0.50	0.03	0.04	
	26	0.55	1	0.004	0	0.50	0	0.96	

Table 6 Stubborn game on Reddit network.

- All chosen nodes are influential in the network
- Node 271 is the only Maxs agent that is not node 0s(Min's agent) neighbor
- Min's agent(node 0) neighbors are is 1, 26, 189, 200, 274, 389, 433, 458

Stubborn Game - NE Opinion Properties

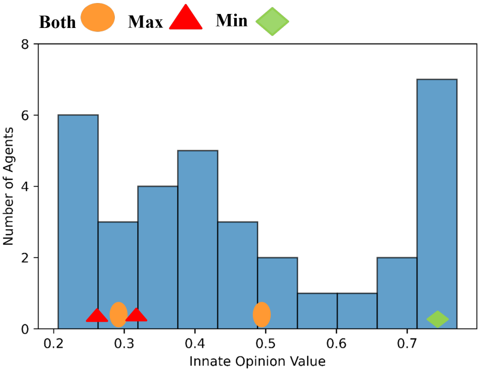


Figure 6 KC opinion distribution

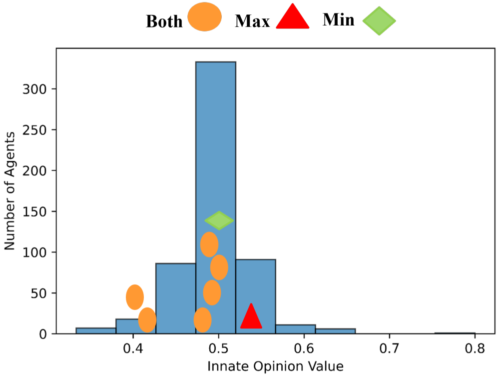
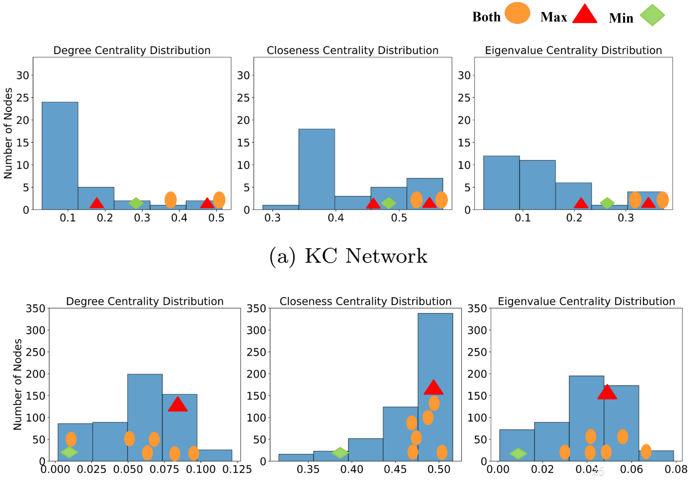


Figure 7 Reddit opinion distribution

Stubborn Agent - NE Centrality Properties



(a) KC Network

(b) Reddit Network

Figure 8 Nodes centrality distribution in network

- KC
  - Nodes with top 4 eigenvalue centrality(high influential) are chosen by players.
- Reddit
  - As the network gets large, players do NOT choose the most influential nodes
  - Both players try to mitigate each other's effects by choosing neighbors of the opponent's agents.
  - Minimizer's main agent 0 is less influential than other selected agents.

Conclusion

- We have empirically studied polarization games, to the best our knowledge, this is the first work to study polarization games.
- Analyzed the functional properties of zero-sum games including the optimal responses for players and properties of Nash equilibria.
- Real network simulations match theoretical results on player behaviors in zero-sum games.
- Empirical modeling of polarization dynamics over networks and the analysis of equilibrium strategies for external players aiming to maximize or minimize polarization.
- Extension: Various network dynamic modes, such as static network, set information sources in the network

Questions?

- Thanks for listening!

[1] W. W. Zachary, “An information flow model for conflict and fission in small groups,” Journal of anthropological research, vol. 33, no. 4, pp. 452–473, 1977.

[2] A. De, S. Bhattacharya, P. Bhattacharya, N. Ganguly, and S. Chakrabarti, “Learning a linear influence model from transient opinion dynamics,” in Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, 2014, pp. 401–410.