Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 50-1 (2017) 11896–11901

Opinion evolution in time-varying social influence networks with prejudiced agents

Anton V. Proskurnikov[1,2,3] **Roberto Tempo**[4] **Ming Cao**[5] **Noah E. Friedkin**[6]

[1] *Delft Center for Systems and Control (DCSC), Delft University of Technology, The Netherlands*
[2] *Institute for Problems of Mechanical Engineering (IPME RAS), St. Petersburg, Russia*
[3] *ITMO University, St. Petersburg, Russia*
[4] *CNR-IEIIT, Politecnico di Torino, Torino, Italy*
[5] *Engineering and Technology Institute (ENTEG), University of Groningen, The Netherlands*
[6] *Center for Control, Dynamical Systems and Computation, University of California Santa Barbara, Santa Barbara, USA*

Abstract: Investigation of social influence dynamics requires mathematical models that are "simple" enough to admit rigorous analysis, and yet sufficiently "rich" to capture salient features of social groups. Thus, the mechanism of iterative opinion pooling from (DeGroot, 1974), which can explain the generation of consensus, was elaborated in (Friedkin and Johnsen, 1999) to take into account individuals' ongoing attachments to their initial opinions, or *prejudices*. The "anchorage" of individuals to their prejudices may disable reaching consensus and cause disagreement in a social influence network. Further elaboration of this model may be achieved by relaxing its restrictive assumption of a time-invariant influence network. During opinion dynamics on an issue, arcs of interpersonal influence may be added or subtracted from the network, and the influence weights assigned by an individual to his/her neighbors may alter. In this paper, we establish new important properties of the (Friedkin and Johnsen, 1999) opinion formation model, and also examine its extension to time-varying social influence networks.

1. INTRODUCTION

During the past decades, there has been a substantial growth of interest in dynamics of social influence networks and opinion formation mechanisms in them. In contrast to the recent research emphasis on multi-agent consensus and coordination, models are being advanced that explain observed behaviors of social groups such as disagreement, polarization, and conflict (Friedkin, 2015; Proskurnikov and Tempo, 2017). An explanatory network science is advancing on the structural properties of social networks (Wasserman and Faust, 1994; Easley and Kleinberg, 2010) and some special dynamical processes over these networks, e.g. epidemic spread (Newman, 2003). At the same time, there is a growing recognition that systems and control theories may substantially broaden the scope of our understanding of the definitional problem of sociology—the coordination and control of social systems (Friedkin, 2015).

System-theoretic examination of social dynamics requires mathematical models that are capable of capturing the complex behavior of a social group yet simple enough to be rigorously examined. In this paper, we deal with one such model, proposed by Friedkin and Johnsen (Friedkin and Johnsen, 1999, 2011; Friedkin, 2015) and henceforth referred to as the FJ model. The FJ model extends the idea of iterative "opinion pooling" (DeGroot, 1974) by assuming that some agents are *prejudiced*. These agents have some level of "anchorage" on their initial opinions (*prejudices*) and factor them into any iteration of their opinions. Similar to continuous-time clustering protocols with "informed" leaders (Xia and Cao, 2011), the heterogeneity of the prejudices and its linkage to individuals' susceptibilities to interpersonal influence may lead to persistent disagreement of opinions and outcomes such as polarization and clustering. With the FJ model, the clustering of opinions does not require the existence of repulsive couplings, or "negative ties" among individuals (Fläche and Macy, 2011; Altafini, 2013; Proskurnikov et al., 2016a; Xia et al., 2016) whose ubiquity in interpersonal interactions is still waiting for supporting experimental evidence (Takács et al., 2016). Unlike models with discrete opinions (Castellano et al., 2009) and bounded confidence models (Hegselmann and Krause, 2002; Weisbuch et al., 2005; Blondel et al., 2009), the FJ model describes the opinion evolution by *linear* discrete-time equations, and is thus much simpler for mathematical analysis. At the same time, the FJ model has been confirmed by experiments with real social groups (Friedkin and Johnsen, 2011; Friedkin et al., 2016a). The FJ model is closely

⋆ Partial funding was provided by NWO (vidi-438730), ERC (grant ERC-StG-307207), CNR International Joint Lab COOPS, Russian Federation President's Grant MD-6325.2016.8 and RFBR, grants 17-08-01728, 17-08-00715 and 17-08-01266. Theorem 2 is obtained under sole support of Russian Science Foundation grant 14-29-00142. E-mails: anton.p.1982@ieee.org, m.cao@rug.nl, friedkin@soc.ucsb.edu

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2017.08.1424

related to the PageRank algorithm (Friedkin and Johnsen, 2014; Proskurnikov et al., 2016b) and has been given some elegant game-theoretic and electric interpretations (Bindel et al., 2011; Ghaderi and Srikant, 2014; Frasca et al., 2015). In the recent works (Parsegov et al., 2017; Proskurnikov and Tempo, 2017) necessary and sufficient conditions for the stability of the FJ model has been established; these conditions also provide convergence "on average" of its decentralized gossip-based counterpart (Frasca et al., 2013; Ravazzi et al., 2015; Frasca et al., 2015). A multidimensional extension of the FJ model has been used to describe the evolution of *belief systems* (Parsegov et al., 2017; Friedkin et al., 2016b), representing invidiuals' positions on several *mutually dependent* issues.

In this paper, we further develop the mathematical theory of the FJ model, obtaining explicit estimates for its convergence speed. We also examine an extension of the classical FJ model, describing a natural *time-varying* social influence process. Such an extension is important since during opinion dynamics on an issue, arcs of interpersonal influence may be added or subtracted from the network, and the influence weights assigned by an individual to his/her neighbors may alter. An example of such an evolution is the dynamics of individuals' *reflected appraisals* (Jia et al., 2015; Friedkin et al., 2016a; Chen et al., 2016).

2. PRELIMINARIES AND NOTATION

We denote matrices with capital letters $A = (a_{ij})$, using lower case letters for their scalar entries and vectors. The symbol $\mathbf{1}_n$ denotes the column vector of ones $(1, 1, \dots, 1)^\top \in \mathbb{R}^n$, and I_n is the identity $n \times n$ matrix. For two vectors $x, y \in \mathbb{R}^n$ we write $x \le y$ if $x_i \le y_i \forall i$. The spectral radius of a square matrix A is denoted by $\rho(A)$, the matrix is *Schur* stable if $\rho(A) < 1$. A nonnegative matrix A is *substochastic* if $\sum_j a_{ij} \le 1$ for any i. Any such matrix has $\rho(A) \le 1$ due to the Gershgorin disk theorem (Horn and Johnson, 1985). A substochastic matrix A is *stochastic* if $\sum_j a_{ij} = 1 \forall i$; when A is sized $n \times n$, the stochasticity implies that $A\mathbf{1}_n = \mathbf{1}_n$ and $\rho(A) = 1$.

A (weighted directed) *graph* is a triple $\mathcal{G} = (\mathcal{V}, \mathcal{E}, W)$, where $\mathcal{V} = \{v_1, \dots, v_n\}$ stands for the set of *nodes*, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of *arcs*, and W is a (weighted) $n \times n$ *adjacency matrix*, i.e. $w_{ij} > 0$ when $(i, j) \in \mathcal{E}$ and otherwise $w_{ij} = 0$. Henceforth we assume that $\mathcal{V} = \{1, 2, \dots, n\}$ and thus the graph $\mathcal{G} = \mathcal{G}(W)$ is uniquely defined by its adjacency matrix W. We denote an arc $(i, j) \in \mathcal{E}$ by $i \mapsto j$ and call the value w_{ij} its *weight*. A chain of arcs $i_0 \mapsto i_1 \mapsto \dots \mapsto i_{r-1} \mapsto i_r$ is a *walk* of length r from node i_0 to node i_r.

3. THE FRIEDKIN-JOHNSEN MODEL

The FJ model describes a network of *social influence* (Friedkin and Johnsen, 2011), consisting of n individuals, or social *agents* indexed 1 through n. The agents opinions are represented by scalars $x_i \in \mathbb{R}$, constituting the vector of opinions $x = (x_1, \dots, x_n)^\top$. The process of social influence is described by two matrices: a stochastic matrix of *interpersonal influences* $W \in \mathbb{R}^{n \times n}$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_{11}, \dots, \lambda_{nn})$ of individual *susceptibilities* $\lambda_{ii} \in [0; 1]$ to the interpersonal influence. At each step, the vector of opinions changes as follows

$$x(k+1) = \Lambda W x(k) + (I_n - \Lambda)u, \quad k = 0, 1, \dots. \qquad (1)$$

The elements u_i of the constant vector u stand for the agents' *prejudices*; the original FJ model (Friedkin and Johnsen, 1999; Friedkin, 2015) assumed that $u_i = x_i(0)$.

In the special case where $\Lambda = I_n$ the model (1) reduces to DeGroot's iterative "opinion pooling" (DeGroot, 1974), providing a discrete-time *consensus algorithm* (Ren and Beard, 2008). At each step, an agent sets its new opinion to be the convex combination of its own and others' opinions

$$x_i(k+1) = \sum_{j=1}^n w_{ij} x_j(k) \, \forall i \Longleftrightarrow x(k+1) = W x(k). \qquad (2)$$

The influence weight w_{ij} shows the contribution of jth opinion on each stage to the ith opinion on the next stage.

The FJ model (1) also employs the mechanism of convex combination, allowing some agents to be prejudiced. If $\lambda_{ii} < 1$ then agent i is "attached" to its prejudice u_i and factors it into any opinion iteration, replacing (2) by

$$x_i(k+1) = \lambda_{ii} \sum_{j=1}^n w_{ij} x_j(k) + (1 - \lambda_{ii}) u_i \, \forall i. \qquad (3)$$

When $\lambda_{ii} = 1$, the ith agent's opinion is formed by the DeGroot mechanism (2), otherwise its prejudice influences each stage of the opinion iteration. Agent i with $\lambda_{ii} = 0$ is "totally prejudiced" and its opinion is static $x_i(k) \equiv u_i$.

Under the assumption $u_i = x_i(0)$, adopted in the FJ model, any agent with $w_{ii} = 1$ (and thus $w_{ij} = 0 \forall j \ne i$) retains its opinion constant $x_i(k) = u_i$ independent of λ_{ii}, and one may suppose, without loss of generality, that

$$w_{ii} = 1 \Longleftrightarrow \lambda_{ii} = 0. \qquad (4)$$

In the original model from (Friedkin and Johnsen, 1999) an even stronger *coupling condition* $w_{ii} = 1 - \lambda_{ii} \forall i$ was adopted for parsimony in the model's empirical applications. In this paper, we do not assume this condition to hold, so Λ and W are independent except for the nondegeneracy condition (4). Notice that each FJ model corresponds to the substochastic matrix $A = \Lambda W$; for the models satisfying (4) this correspondence is one-to-one. A substochastic matrix A is decomposed as $A = \Lambda W$, where

$$\lambda_{ii} = \sum_j a_{ij} \quad \text{and} \quad w_{ij} = \begin{cases} a_{ij}/\lambda_{ii}, & \lambda_{ii} > 0, \\ 1, & i = j \text{ and } \lambda_{ii} = 0, \\ 0, & i \ne j \text{ and } \lambda_{ii} = 0. \end{cases}$$

The stability criteria for FJ models may thus be reformulated for substochastic matrices, and vice versa.

For us it will be convenient to discard the standard assumption $x(0) = u$ and consider u as some constant external "input", independent of the initial opinion [1] $x(0)$.

A central question concerned with the FJ dynamics (1) is its *convergence* of opinion vectors to a finite limit

$$x^\infty = \lim_{k \to \infty} x(k). \qquad (5)$$

[1] Individuals prejudices may be explained (Friedkin and Johnsen, 1999) by the system "history", e.g. the effect of some exogenous factors, which influenced the community in the past. This motivates to introduce the explicit relation between the prejudice and initial condition of the social system $u = x(0)$. However, the prejudices can also be some non-trivial functions of the initial conditions $u = u(x(0))$ or be caused by external factors that are not related to the system's history, e.g. some information spread in social media.