SIAM J. MATRIX ANAL. APPL.
Vol. 31, No. 4, pp. 1812–1834
© 2010 Society for Industrial and Applied Mathematics

ITERATIVE FILTERING IN REPUTATION SYSTEMS*

CRISTOBALD DE KERCHOVE† AND PAUL VAN DOOREN†

Abstract. We present a class of voting systems that we call "iterative filtering" systems. These systems are based on an iterative method that assigns a reputation to $n + m$ items, n objects, and m raters, applying some filter to the votes. Each rater evaluates a subset of objects leading to an $n \times m$ rating matrix with a given sparsity pattern. From this rating matrix a formula is defined for the reputation of raters and objects. We propose a natural and intuitive nonlinear formula and also provide an iterative algorithm that linearly converges to the unique vector of reputations. In contrast to classical outlier detection, no evaluation is discarded in this method, but each one is taken into account with different weights for the reputations of the objects. The complexity of one iteration step is linear in the number of evaluations, making our algorithm efficient for large data sets. Experiments show good robustness of the reputation of the objects against cheaters and spammers and good detection properties of cheaters and spammers.

Key words. reputation systems, nonlinear dynamical systems, iterative filtering, voting, optimization

AMS subject classifications. 34D05, 37C25, 70K20

DOI. 10.1137/090748196

Introduction: Motivations and challenges in voting systems.

Short review of voting systems. Many measures of reputation have been proposed over the years under the names of reputation, voting, ranking, or trust systems, and they deal with various contexts ranging from the classification of football teams to the reliability of each individual in peer-to-peer systems. Surprisingly enough, the most used method for reputation on the Web amounts simply to averaging the votes. In that case, the reputation is, for instance, the average of scores represented by five stars on YouTube, or the percentage of positive transactions on eBay. Therefore such a method evenly trusts each rater of the system. Besides this method, many other algorithms exploit the structure of networks generated by the votes: raters and evaluated items are nodes connected by votes. A great many of these methods use efficient eigenvector-based techniques or trust propagation over the network to obtain the reputation of every node [2, 3, 4, 5, 6, 7, 8]. They can be interpreted as a distribution of some reputation flow over the network where reputations satisfy some transitivity: you have a high reputation if you have several incoming links coming from nodes with a high reputation. The average methods, the eigenvector-based techniques and trust propagation, may suffer from noise in the data and bias from dishonest raters. For this reason, they are sometimes accompanied by statistical methods for spam detection [9, 10], as in the context of web pages trying to boost their PageRank scores by adding artificial incoming links [11, 12]. Detected spam is then simply removed from

*Received by the editors January 27, 2009; accepted for publication (in revised form) by I. Ipsen December 3, 2009; published electronically March 31, 2010. This work has been supported by the Concerted Research Action (ARC) "Large Graphs and Networks" from the "Direction de la recherche scientifique—Communauté française de Belgique," and by the Belgian Programme on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office.

http://www.siam.org/journals/simax/31-4/74819.html

†Department of Applied Mathematics, Ecole Polytechnique de Louvain, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium (c.dekerchove@uclouvain.be, paul.vandooren@uclouvain.be).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

the data. Therefore, the three main strategies for voting systems are simple methods averaging votes where raters are evenly trusted, eigenvector-based techniques and trust propagation, where reputations depend directly on reputations of the neighbors, and, finally, statistical measures to classify and possibly remove some of the items.

Iterative filtering systems. We will propose a method drawn largely from the algorithm of Laureti et al. in [14] and Yu et al. in [15], where we compare the evaluations of the raters to deduce some weights of trust for these votes and then update the reputations of each item accordingly. Obviously the choice of a specific reputation system, with specific interpretations of the votes, depends on subjective properties that we just accept. For example, in the average method, we agree that every rater is taken into account in the same manner. In the PageRank algorithm, we accept that a random walk over the network is a good model of real navigation for a web surfer. In trust propagation over networks, we accept the transitivity of trust: if A trusts B and B trusts C, then A will trust C.

Concerning the iterative filtering (IF) systems, we will make the following assumption: *raters whose opinions often differ from those of other raters are assigned less weight.* We label this the *IF*-property and will formally define it later on. This property is at the heart of the filtering process and implies that all votes are taken into account, but with a continuous validation scale, in contrast with the explicit deletion of outliers. Moreover, the weight of each rater depends on the distance between his votes and the reputation of the objects he evaluates: typically weights of random raters and outliers decrease during the iterative filtering. The main criticism one can have about the *IF*-property is that it discriminates against "marginal" evaluators, i.e., raters who vote differently from the average opinion for many objects. However, IF systems may have several basins of attraction, each corresponding to a group of people with a coherent opinion.

Dynamical data. Votes, raters, and objects can appear, disappear, or change, making the system dynamical. This is, for example, the case when we consider a stream of news as in [13]: news sources and articles are ranked according to their publication over time. Nowadays, most sites driven by raters involve dynamical opinions. For instance, blogs, the site Digg, and the site Flickr are good places to exchange and discuss ideas, remarks, and votes about various topics ranging from political elections to photos and videos. We will see that IF systems allow evolving rating matrices and then provide time-varying reputations.

Structure of the paper. Section 1 introduces the definitions and the properties of IF systems. Then section 2 discusses the choice of the discriminant function that allows us to penalize the raters differently according to their votes. Section 3 focuses on our choice of discriminant function (that is, affine) and analyzes the convergence of IF systems with this function. Section 4 extends some results to sparse rating matrices and dynamical data. Section 5 presents simulations on two real data sets that deal with the effect of malicious users (not analyzed in the previous section). Section 6 concludes and gives some perspectives for further research.

1. Iterative filtering systems: Definitions and properties. For the sake of clarity, we first consider the case where the votes are fixed, i.e., the rating matrix does not change over time, and all objects are evaluated by all raters, i.e., the rating matrix is full. The dynamical case and the sparsity pattern for the rating matrix will be analyzed in section 4. With these assumptions, we present the main properties of IF systems and then we restrict ourselves to the natural case of quadratic IF systems,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

where the reputations are given by a linear combination of the votes and the weights of the raters are based on the Euclidean distance between the reputations and the votes.

1.1. General notation. Let $X \in \mathbb{R}^{n\times m}$ be the rating matrix, $\boldsymbol{r} \in \mathbb{R}^n$ be the reputation vector of the objects, and $\boldsymbol{w} \in \mathbb{R}^m$ be the weight vector of the raters. The entry X_{ij} is the vote to object i given by rater j, and the vector $\boldsymbol{x}_j$, the jth column of X, represents the votes of rater j:

$$X = [\boldsymbol{x}_1 \ \dots \ \boldsymbol{x}_m] .$$

The bipartite graph formed by the objects, the raters, and their votes is represented by the $n \times m$ adjacency matrix A; i.e., $A_{ij} = 1$ if object i is evaluated by rater j, and 0 otherwise. For the sake of simplicity, we assume in this section that every object has been evaluated by all raters $A_{ij} = 1$ for all i, j. The general case where the bipartite graph is not necessarily complete will be handled later.

The belief divergence d_j of rater j is the normalized distance between his votes and the reputation vector $\boldsymbol{r}$,

$$\boldsymbol{d} = \frac{1}{n} \begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{r}\|_2^2 \\ \vdots \\ \|\boldsymbol{x}_m - \boldsymbol{r}\|_2^2 \end{pmatrix} . \tag{1.1}$$

Therefore equations (1.1) are quadratic equations in $\boldsymbol{r}$ and give an estimate of the variances of the votes for every rater compared to a given reputation vector $\boldsymbol{r}$.

1.2. Definition of IF systems. Before introducing IF systems, we define the two basic functions of these systems:

$$\text{(1) the reputation function} \qquad F : \mathbb{R}^m \to \mathbb{R}^n : F(\boldsymbol{w}) = \boldsymbol{r},$$

which gives the reputation vector depending on the weights of the raters and implicitly on the rating matrix X;

$$\text{(2) the filtering function} \qquad G : \mathbb{R}^n \to \mathbb{R}^m_{\geq 0} : G(\boldsymbol{r}) = \boldsymbol{w},$$

which gives the weight vector for the raters depending on the belief divergence $\boldsymbol{d}$ of each rater defined in (1.1).

We formalize the so-called *IF-property* described in the introduction that assumes that raters diverging often from the opinion of other raters are assigned less weight. We will make the reasonable assumption that raters with identical belief divergence receive equal weights. Hence, we can write

$$G(\boldsymbol{r}) = \begin{bmatrix} g(d_1) \\ \vdots \\ g(d_m) \end{bmatrix} . \tag{1.2}$$

We call the scalar function g the discriminant function associated with G. Equation (1.2) indicates that every rater has the same discriminant function g, but we could also consider personalized functions g_j penalizing raters to different degrees.

A filtering function G satisfies the *IF*-property if its associated discriminant function $g : \mathbb{R} \to \mathbb{R}$ is nonnegative and monotonically decreasing. Therefore, the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

IF-property merely implies that a decrease in belief divergence d_j for any rater j corresponds to a larger (or equal) weight w_j. Laureti et al. [14] and Yu et al. [15] introduced the case where the discriminant function is given by $g(d) = d^{-k}$ (with $k \geq 0$). We propose $g(d) = 1 - k\,d$ as discriminant function with some condition on k to avoid negative weights. These two choices will be discussed later and compared in a small example.

1.3. Quadratic IF systems. We define the *quadratic IF systems* by fixing the reputation function $F(\boldsymbol{w})$ that is naturally given by taking the weighted average of the votes:

$$F(\boldsymbol{w}) = X\boldsymbol{w} / \sum_{j=1}^{m} w_j \tag{1.3}$$

for any nonnegative vector $\boldsymbol{w}$. If we assume that the votes are in $[a, b]$, we then have that the reputation vector $\boldsymbol{r}$ belongs to the hypercube

$$\mathcal{H} := [a, b]^n \tag{1.4}$$

and, more precisely, that it belongs to the convex hull $\mathcal{P} \subseteq \mathcal{H}$ for the set of points $\{\boldsymbol{x}_j : j = 1, \ldots, p\}$; that is,

$$\mathcal{P} = \left\{ \boldsymbol{r} \in \mathbb{R}^n \,|\, \boldsymbol{r} = \sum_{j=1}^{m} w_j \boldsymbol{x}_j \text{ with } \sum_{j=1}^{m} w_j = 1 \text{ and } w_j \geq 0 \right\}. \tag{1.5}$$

DEFINITION 1.1. *Quadratic IF systems are systems of equations in the reputations $\boldsymbol{r}^t$ of the objects and the weights $\boldsymbol{w}^t$ of the raters that evolve over discrete time t according to the rating matrix X:*

$$\boldsymbol{r}^{t+1} = F(\boldsymbol{w}^t) = X\boldsymbol{w}^t / \sum_{j=1}^{m} w_j, \tag{1.6}$$

$$\boldsymbol{w}^{t+1} = G(\boldsymbol{r}^{t+1}), \tag{1.7}$$

starting with equal weights $\boldsymbol{w}^0 = \mathbf{1}$.

Then our method uses the affine function $g(d) = 1 - k\,d$. This leads to the following definition.

DEFINITION 1.2. *Our method is a quadratic IF system with affine discriminant function $g(d) = 1 - k\,d$ for some positive k with*

$$\boldsymbol{r}^{t+1} = X\boldsymbol{w}^t / \sum_{j=1}^{m} w_j, \tag{1.8}$$

$$\boldsymbol{w}^{t+1} = \mathbf{1} - k\,\frac{1}{n} \begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{r}^{t+1}\|_2^2 \\ \vdots \\ \|\boldsymbol{x}_m - \boldsymbol{r}^{t+1}\|_2^2 \end{pmatrix}, \tag{1.9}$$

starting with equal weights $\boldsymbol{w}^0 = \mathbf{1}$.

We will show in the next subsection that quadratic IF systems correspond to taking the steepest descent direction and minimizing a particular energy function.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

1.4. Energy functions for quadratic IF systems. By using the trivial relation $\sum_{j=1}^{m} w_j = \mathbf{1}^T \boldsymbol{w}$, we can reformulate one iteration step of $\boldsymbol{r}^t$ for a quadratic IF system as $\boldsymbol{r}^{t+1}(\mathbf{1}^T \boldsymbol{w}^t) = X\boldsymbol{w}^t$; hence a fixed point $\boldsymbol{r}^*$ (if it exists) is given by quadratic equations in $\boldsymbol{r}^*$ and $\boldsymbol{w}^*$,

$$\boldsymbol{r}^*(\mathbf{1}^T \boldsymbol{w}^*) = X\boldsymbol{w}^*, \tag{1.10}$$

where $\boldsymbol{w}^* = G(\boldsymbol{r}^*)$. Hence, for any integrable function G, the fixed points in (1.10) are the roots of the function

$$D(\boldsymbol{r}) := \frac{2}{n}(\boldsymbol{r}\mathbf{1}^T - X) \cdot G(\boldsymbol{r}), \tag{1.11}$$

which is the gradient of some scalar function introduced in the next theorem and labeled the energy function E.

THEOREM 1.3. *The fixed points of quadratic IF systems with integrable discriminant function g are the stationary points of the energy function*

$$E(\boldsymbol{r}) = \sum_{j=1}^{m} \int_0^{d_j(\boldsymbol{r})} g(u)\, du + c, \tag{1.12}$$

where d_j is the belief divergence of rater j that depends on $\boldsymbol{r}$, and $c \in \mathbb{R}$ is a constant. Moreover, one iteration step in quadratic IF systems corresponds to a steepest descent direction with a particular step size

$$\boldsymbol{r}^{t+1} = \boldsymbol{r}^t - \alpha^t \nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t), \quad \text{with } \alpha^t = \frac{n}{2\sum_{j=1}^{m} w_j}. \tag{1.13}$$

Proof. We have $\nabla_{\boldsymbol{r}} E(\boldsymbol{r}) = \nabla_{\boldsymbol{r}} \boldsymbol{d}^T \cdot \nabla_{\boldsymbol{d}} E(\boldsymbol{r})$ with

$$\nabla_{\boldsymbol{r}} \boldsymbol{d}^T = -\frac{2}{n}(X - \boldsymbol{r}\mathbf{1}^T) \quad \text{and} \quad \nabla_{\boldsymbol{d}} E(\boldsymbol{r}) = G(\boldsymbol{r}).$$

Therefore a stationary point $\boldsymbol{r}^*$ in E satisfies

$$\begin{aligned} -\frac{2}{n}(X - \boldsymbol{r}^*\mathbf{1}^T) \cdot G(\boldsymbol{r}^*) &= 0, \\ (X - \boldsymbol{r}^*\mathbf{1}^T)\boldsymbol{w}^* &= 0, \\ \boldsymbol{r}^*(\mathbf{1}^T \boldsymbol{w}^*) &= X\boldsymbol{w}^*, \end{aligned}$$

which corresponds to the fixed point equation given in (1.10).

We also have $\nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t) = \nabla_{\boldsymbol{r}} (\boldsymbol{d}(\boldsymbol{r}^t))^T \cdot \nabla_{\boldsymbol{d}} E(\boldsymbol{r}^t)$ with

$$\nabla_{\boldsymbol{r}} (\boldsymbol{d}(\boldsymbol{r}^t))^T = -\frac{2}{n}(X - \boldsymbol{r}^t\mathbf{1}^T) \quad \text{and} \quad \nabla_{\boldsymbol{d}} E(\boldsymbol{r}^t) = G(\boldsymbol{r}^t) = \boldsymbol{w}^t.$$

Therefore

$$\begin{aligned} \nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t) &= -\frac{2}{n}(X - \boldsymbol{r}^t\mathbf{1}^T)\boldsymbol{w}^t \\ &= -\frac{2}{n}(\mathbf{1}^T \boldsymbol{w}^t)(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t) = -\frac{1}{\alpha^t}(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t). \end{aligned} \tag{1.14}$$

∎

The stable fixed points of quadratic IF systems minimize the sum of the integrals $\int_0^{d_j} g(u)\, du$ on $j = 1, \ldots, m$, meaning that they minimize the sum of surfaces below g in the intervals $[0, d_j]$ for $j = 1, \ldots, m$. For example, when g is constant, the weights are always equal and the unique fixed point is given by the average of the votes minimizing $\|\boldsymbol{d}\|_1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

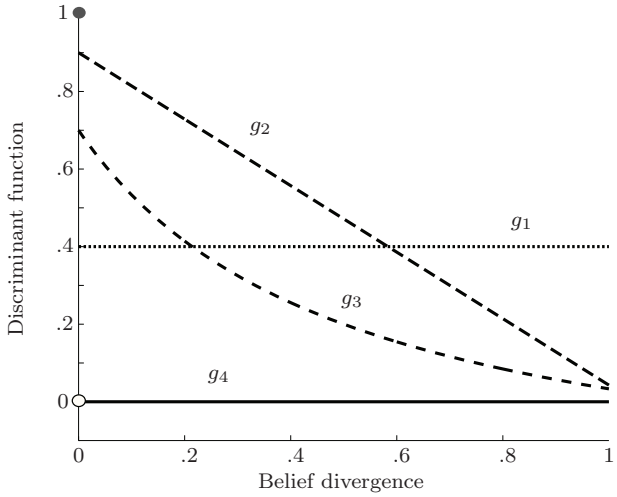

FIG. 2.1. (g_1) *The discriminant function is constant, and weights are independent from the belief divergence;* (g_2, g_3) *intermediate cases; and* (g_4) *weights are null except at the origin.*

2. The discriminant function. We discuss in this section the different properties of quadratic IF systems depending on their discriminant function g.

2.1. The compromise. Two extreme cases are shown in Figure 2.1: the constant function g_1 leads to equal weights and consequently to merely averaging the votes for the reputations. In that case, the belief divergence is not taken into account for the calculation of the weights, and the solution is unique. On the other hand, the function g_4 that gives positive weights only for the raters with zero belief divergence leads to quadratic IF systems with only m reputations. Therefore, the fixed points correspond to very local opinions, where the weight of rater j is maximal, while the weights of other raters are minimal.

Clearly, we are interested in intermediate cases with the following compromise: the discriminant function must be sufficiently discriminating to penalize the outliers, but it has to avoid meaningless local solutions.

2.2. One versus several solutions. The number of stable fixed points is determined by the rating matrix X and the discriminant function g. The former gives the distribution of the votes and can have clusters of opinions in the data. The latter can be defined such that the cost function always has a unique minimum, or such that several clusters lead to several minima.

When the function g guarantees the uniqueness of the solution, we avoid the choice between different solutions and we also simplify the dynamics of IF systems since multiple stable points imply the existence of unstable points or bifurcations.

Let us remark that dynamical IF systems with possibly several minimizers mean that we need to track several trajectories during the iterations. Every fixed point will have a basin of attraction depending on its iteration function. We will then need to provide methods that avoid reproducing identical solutions.

2.3. The candidates. Let us introduce three choices of function g having interesting properties:

$$g(d) = d^{-k}, \tag{2.1}$$

$$g(d) = e^{-k\,d}, \tag{2.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$g(d) = 1 - k\,d. \tag{2.3}$$

All discriminant functions g are positive and decrease with d for positive k and therefore satisfy the IF-property. However k must be small enough to keep g positive in (2.3) and hence to avoid negative weights. We will see that this condition is sufficient for (2.3) to guarantee the uniqueness of the minimum of the energy function E defined in (1.12). However, the condition of uniqueness is different and not established for (2.1), (2.2). When $k = 0$, we recover the average method, where the weights are equal. For increasing values of k, the functions g become more and more penalizing. Let us analyze in more detail these three candidates which we name the inverse, exponential, and affine functions.

Inverse function. The first definition was proposed in [14, 15] with $k \geq 0$ (in particular, $k = 1$ and $k = 1/2$). The choice of function g comes from the maximization of some multivariate density function representing the distribution of the votes. There d_j is taken as an estimate of the variance σ_j^2 of rater j's votes. In most cases, the method converges to a unique stable fixed point. However, if there exist several divergent opinions among the raters (that are sufficiently far away from each other), then more than one stable fixed point may appear. In addition, unstable fixed points can make the iterations sensitive to initial conditions, as we will see in a small example.

Exponential function. The function g in (2.2) leads to another quadratic IF system based on a similar argument. The probability density function f for the votes $\boldsymbol{x}_j$ of rater j with given $\boldsymbol{r}$ and the positive scalar σ is

$$f(\boldsymbol{x}_j|\boldsymbol{r},\sigma) = \prod_i \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(X_{ij}-r_i)^2}{\sigma^2}}. \tag{2.4}$$

This time, we assume that the raters have the same variance σ for their votes. Assuming the same variance for all raters leads to some limitations in the model. More complex models have been developed in [14, 15] to take into account different variances and a bias for each rater. However, the model can be made more complex in further studies (e.g., different discriminant functions for the raters).

We recover the quadratic IF system with $g(d) = e^{-k\,d}$ by considering the iterations on the reputations, $\boldsymbol{r}^{t+1} = X\frac{\boldsymbol{w}^t}{\mathbf{1}^T\boldsymbol{w}^t}$, and those on the weights (up to a scaling factor),

$$\boldsymbol{w}^{t+1} = \begin{bmatrix} f(\boldsymbol{x}_1|\boldsymbol{r}^{t+1},\sigma) \\ \vdots \\ f(\boldsymbol{x}_m|\boldsymbol{r}^{t+1},\sigma) \end{bmatrix} = \left(\sqrt{2\pi}\sigma\right)^{-n} \begin{bmatrix} e^{-\frac{n}{\sigma^2}d_1^{t+1}} \\ \vdots \\ e^{-\frac{n}{\sigma^2}d_m^{t+1}} \end{bmatrix}.$$

The parameter k is then proportional to the inverse of the variance, that is, n/σ^2. Therefore, by assuming a large variance for all votes, we decrease k, and this naturally implies a larger acceptance of divergent opinions, making the function g less discriminating.

Affine function. Function g in (2.3) links the belief divergence $\boldsymbol{d}$ to the weights $\boldsymbol{w}$ in a simple way: by an affine function. It has the advantage of making the analysis tractable with a direct condition of uniqueness on the parameter k. Its associated quadratic IF system also has a statistical interpretation. For this purpose, we use the log-likelihood of the density function in (2.4) for the iterations of the weights. This gives a degree of belief for each rater that was used in [9] but without iterative process. The quadratic IF system with $g(d) = 1 - k\,d$ is recovered by considering the

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

iterations on the reputations $\boldsymbol{r}^{t+1} = X\frac{\boldsymbol{w}^t}{\mathbf{1}^T\boldsymbol{w}^t}$ and those on the weights (up to a scaling factor)

$$\boldsymbol{w}^{t+1} = \log \begin{bmatrix} f(\boldsymbol{x}_1|\boldsymbol{r}^{t+1},\sigma) \\ \vdots \\ f(\boldsymbol{x}_m|\boldsymbol{r}^{t+1},\sigma) \end{bmatrix},$$

where log is applied componentwise and parameter k is proportional to $\frac{2n}{\sigma^2}\log\frac{1}{2\pi\sigma^2}$. The same remark as before can be made about the link between the variance of the votes and the discriminant parameter k.

Let us give in the following table the equivalent energy functions for the four discriminant functions we discussed.

Function g	$\boldsymbol{r}^*$ minimizes
$g(d) = \frac{1}{d}$	$\sum_j \log d_j$
$g(d) = \frac{1}{\sqrt{d}}$	$\sum_j -\sqrt{d_j}$
$g(d) = e^{-k\,d}$	$\sum_j -e^{-k\,d_j}$
$g(d) = 1 - k\,d$	$\sum_j d_j - \frac{k}{2}\sum_j d_j^2$

Let us remark that the function $g(d) = 1 - k\,d$ implies that the system minimizes $\sum_j d_j - \frac{k}{2}\sum_j d_j^2$. The first term $\sum_j d_j$ is minimized by taking the average votes for $\boldsymbol{r}$, and the second term $-\frac{k}{2}\sum_j d_j^2$ is minimized by taking $\boldsymbol{r}$ in $[0,1]^n$. Therefore we have a compromise between the simple average and a solution on the border of the hypercube that diverges from this average. The parameter k strengthens the impact of belief divergences on the weights and makes the solution move away from the average.

Other functions. The list of possible discriminant functions can be long. Among other things, we can also recover a form of the expected maximization (EM) algorithm [16]. The choice of a function will then be motivated by its convergence properties, its sensitivity on the initial conditions, and the number of minima associated with its energy function.

Small example. In order to illustrate some of the quadratic IF systems presented above with different discriminant function g, we show the sequences $(\boldsymbol{r}^t)$ and $(\boldsymbol{w}^t)$ for a given rating matrix X, say in the context of ice skating, and their limit points $\boldsymbol{r}^*$ and $\boldsymbol{w}^*$. We propose the following fictitious rating matrix:

$$X = \begin{bmatrix} 3.3 & 3.4 & 4.9 \\ 4.2 & 4.5 & 2.8 \end{bmatrix},$$

where the two rows correspond to the votes given to two ice skaters and the three columns represent the three raters. The third rater tries to favor the first ice skater. If equal weights are given to the raters, we obtain the average votes: 3.87 for the first skater and 3.83 for the second. We show the different results according to the choice of discriminant function discussed before.

The table below shows the weights and the reputations for several discriminant functions g and the initial vector $\boldsymbol{w}^0 = \mathbf{1}$.

Rows 1–2. The first row gives reputations not far from the average, and the second considers a larger k which is more penalizing. Consequently, the third rater is more severely weighted, and, therefore, the gap between the two skaters increases. Larger k would continue to decrease the third weight. Results do not change if the initial vector $\boldsymbol{w}^0$ is different.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$g(d)$	$(\boldsymbol{w}^*)^T$	$(\boldsymbol{r}^*)^T$
$1-\frac{1}{5}\,d$	$\begin{bmatrix} 1.0 & 1.0 & 0.8 \end{bmatrix}$	$\begin{bmatrix} 3.8 & 3.9 \end{bmatrix}$
$1-\frac{1}{3}\,d$	$\begin{bmatrix} 1.0 & 1.0 & 0.5 \end{bmatrix}$	$\begin{bmatrix} 3.7 & 4.0 \end{bmatrix}$
d^{-1}	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 3.3 & 4.2 \end{bmatrix}$
d^{-1} $(*)$	$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 4.9 & 2.8 \end{bmatrix}$
$d^{-1/2}$	$\begin{bmatrix} 35.2 & 4.8 & .7 \end{bmatrix}$	$\begin{bmatrix} 3.3 & 4.2 \end{bmatrix}$
e^{-d}	$\begin{bmatrix} 1.0 & 1.0 & 0.1 \end{bmatrix}$	$\begin{bmatrix} 3.4 & 4.2 \end{bmatrix}$

Rows 3–4. The iterations are numerically unstable because the belief divergence d_1 of rater 1 tends to zero, making his weight infinite. Disregarding this instability, the normalized weights tend to $[1\,0\,0]^T$, and therefore the convergent reputations vector corresponds to the first column of X. However, the final reputations and weights depend on the initial vector $\boldsymbol{w}^0$. The fourth row $(*)$ considers $\boldsymbol{w}^0 = [1\,1\,4]^T$ instead of equal weights. Then the third rater gains all the confidence, and the final reputations are given by his votes: the third column of X.

Row 5. In this case, the initial point does not matter, provided that we avoid a zero belief divergence. We see that this function almost disqualifies the third rater: it makes a large difference between the first two raters and is more penalizing than $g(d) = 1 - \frac{1}{3}\,d$.

Row 6. Finally, we look at the iterations given by (2.2): $g(d) = e^{-k\,d}$. From a certain k on, several final reputations are possible depending on the initial vector $\boldsymbol{w}^0$. Similarly as for $g(d) = d^{-1}$, if $\boldsymbol{w}^0$ favors the third rater, then this rater will have an advantageous weight at the end. For $k = 1$, we avoid multiple reputations and we obtain the results in row 6.

3. An iterative filtering method with affine discriminant function. We look at the quadratic IF system with the discriminant function g defined in (2.3). The energy function in (1.12) for the constant $c = -m/2k$ associated with our method in (1.8), (1.9) is given by

$$E(\boldsymbol{r}) = -\frac{1}{2\,k}\,\boldsymbol{w}^T\boldsymbol{w}, \tag{3.1}$$

where $\boldsymbol{w}$ depends on $\boldsymbol{r}$ according to the function $G(\boldsymbol{r})$; i.e.,

$$\boldsymbol{w} = \mathbf{1} - k\,\frac{1}{n}\begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{r}\|_2^2 \\ \vdots \\ \|\boldsymbol{x}_m - \boldsymbol{r}\|_2^2 \end{pmatrix}.$$

Therefore, the energy function is a fourth-order polynomial in $\boldsymbol{r}$:

$$E(\boldsymbol{r}) = -\frac{1}{2\,k}\sum_{j=1}^{m}\left(1 - k\,\frac{1}{n}\|\boldsymbol{x}_j - \boldsymbol{r}\|_2^2\right)^2. \tag{3.2}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

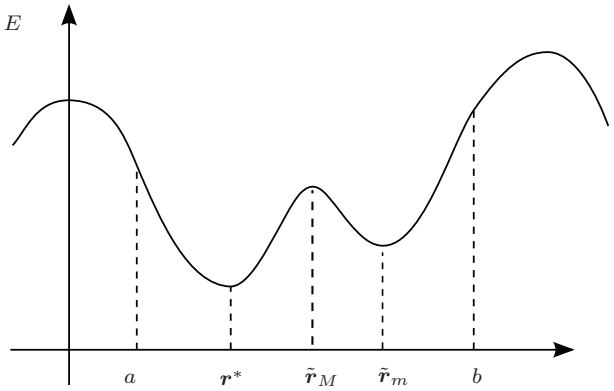

FIG. 3.1. *If $\boldsymbol{r}^*$ is a strict minimum, then another minimum $\tilde{\boldsymbol{r}}_m$ leads to a contradiction. If $\tilde{\boldsymbol{r}}_M$ is a strict maximum, then there would be two other maxima (before a and after b) leading to a contradiction.*

We will see that (under some assumption on k) this energy function decreases with the iteration steps, i.e., the sequence $(E(\boldsymbol{r}^t))$ is monotonically decreasing, and it converges to the unique minimum.

3.1. Uniqueness. The following proposition proves that the stable point of our method is unique. For that purpose, we consider the set of admissible k to guarantee that the weights $\boldsymbol{w} = G(\boldsymbol{r})$ remain positive for every possible reputation vector $\boldsymbol{r}$ in the hypercube $\mathcal{H}$; that is,

$$\mathcal{K} = \left\{ k \in \mathbb{R}_{\geq 0} : \mathbf{1} - k\,\frac{1}{n} \begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{r}\|_2^2 \\ \vdots \\ \|\boldsymbol{x}_m - \boldsymbol{r}\|_2^2 \end{pmatrix} > 0 \text{ for all } \boldsymbol{r} \in \mathcal{H} \right\}. \tag{3.3}$$

The result in the next proposition follows directly from the nature of the energy function E that gives several conditions on the existence of stationary points. These conditions are exposed in the following lemma.

LEMMA 3.1. *Let the function $E : \mathbb{R}^n \to \mathbb{R} : E(\boldsymbol{r}) = z$ be a fourth-order polynomial and let $\mathcal{H}$ be some hypercube in $\mathbb{R}^n$. If*

$$\lim_{\|\boldsymbol{r}\| \to \infty} E(\boldsymbol{r}) = -\infty$$

and the steepest descent direction on the boundary of $\mathcal{H}$ points strictly inside $\mathcal{H}$, then E has a unique stationary point in $\mathcal{H}$ which is a strict minimum.

Proof. Since the steepest descent on the boundary of $\mathcal{H}$ points strictly inside $\mathcal{H}$, there is no stationary point on the boundary, and there is at least one strict minimum in $\text{int}(\mathcal{H})$ that we label $\boldsymbol{r}^*$. Let us prove that the existence of another stationary point leads to a contradiction with the hypothesis.

Let us first assume that $\tilde{\boldsymbol{r}}_m$ is another minimum of E (strict or not). The line passing through the two points $\boldsymbol{r}^*$ and $\tilde{\boldsymbol{r}}_m$ is given by $\ell(y) = \boldsymbol{r}^* + y\,(\tilde{\boldsymbol{r}}_m - \boldsymbol{r}^*)$, and the restriction

$$e(y) := E \circ \ell(y)$$

is a polynomial of degree 4 with two minima in $y = 0$ and $y = 1$. But this is not possible with the hypothesis $\lim_{\|\boldsymbol{r}\| \to \infty} E(\boldsymbol{r}) = -\infty$; see Figure 3.1.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Let us now assume that $\tilde{\boldsymbol{r}}_M$ is a maximum of E (strict or not). The line passing through $\tilde{\boldsymbol{r}}_M$ with the direction $\boldsymbol{e}_1$ is given by $\ell'(y) = \tilde{\boldsymbol{r}}_M + \boldsymbol{e}_1(y - (\tilde{r}_M)_1)$. The restriction

$$e'(y) := E \circ \ell'(y)$$

would be a polynomial of degree 4 with three maxima: one in $]-\infty, 0[$, one at $y = (\tilde{r}_M)_1$, and one in $]1, \infty[$; see Figure 3.1.

Finally, let us assume that $\tilde{\boldsymbol{r}}$ is a saddle point in $\mathcal{H}$. This implies that there is an increasing trajectory starting in $\tilde{\boldsymbol{r}}$ and following the steepest ascent directions. By the condition on the boundary, such a trajectory cannot escape from $\mathcal{H}$. Therefore, it should reach a maximum $\tilde{\boldsymbol{r}}_M$ in $\mathcal{H}$, which is impossible. ∎

PROPOSITION 3.2. *If $k \in \mathcal{K}$, the system in* (1.8), (1.9) *has a unique fixed point $\boldsymbol{r}^*$.*

Proof. For the sake of conciseness, we restrict ourselves to the case where all objects receive at least two different votes. The proof for the general case is similar if we merge the objects with equal votes (the reputations of such objects remain the same).

Since every object has at least two different votes in $[a, b]$ given by two raters with positive weights ($k \in \mathcal{K}$), we have that the sequence[1] $(\boldsymbol{r}^t)$ remains in $\text{int}(\mathcal{H})$. Then by Theorem 1.3, it is sufficient to show that the energy function in (3.2) has a unique stationary point in $\text{int}(\mathcal{H})$.

The steepest descent direction at any point $\boldsymbol{r}^t \in \mathcal{H}$ is given by (1.13),

$$-\nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t) = \frac{1}{\alpha^t}(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t),$$

and since for any point $\boldsymbol{r}^t$ on the boundary of $\mathcal{H}$, the next point $\boldsymbol{r}^{t+1}$ belongs to $\text{int}(\mathcal{H})$ (all weights are strictly positive), the steepest descent direction of E on the boundary of $\mathcal{H}$ points strictly inside $\mathcal{H}$. Therefore, using Lemma 3.1 for E, there is a unique stationary point in $\mathcal{H}$ which is a minimum, and by Theorem 1.3, it is the unique fixed point of the system in (1.8), (1.9). ∎

Figure 3.2 illustrates the stationary points of E when k is taken larger. First, if $k = \sup \mathcal{K}$, then the weights $\boldsymbol{w}$ are nonnegative (rather than positive); therefore maxima and saddle points can appear on the boundary of $\mathcal{H}$ (Figure 3.2(b) is close to this case). Therefore iterations have to avoid these unstable points. Second, if k is strictly larger than $\sup \mathcal{K}$, then maxima can appear inside $\mathcal{H}$; see Figure 3.2(c). Moreover, the existence of a minimum is not guaranteed anymore. However, if it exists, it remains unique, and it can be shown that its basin of attraction contains an open neighborhood of that point. This choice of larger k is discussed at the end of the section.

3.2. The search method. Our method represents the naive method for reaching the minimum of the energy function E in $\mathcal{H}$. We prove that it converges by taking the steepest descent direction at every iteration step.

Before introducing the theorem that proves that the energy function decreases after every iteration step, we present the following lemma that is useful for that theorem. It will also be used later when we will consider a sparsity pattern in the rating matrix. We will use $M \circ M = M^{\circ 2}$ as componentwise product.

LEMMA 3.3. *Given two matrices $M \in \mathbb{R}^{n \times m}$ and $A \in \{0, 1\}^{n \times m}$ such that $M \circ A = M$, we have*

$$[M^T - A^T \circ \mathbf{1}\boldsymbol{c}^T]^{\circ 2}\mathbf{1} = [M^T]^{\circ 2}\mathbf{1} - 2M^T\boldsymbol{c} + A^T\boldsymbol{c}^{\circ 2}.$$

[1] To emphasize the importance of the order, we are not using the standard notation $\{\boldsymbol{r}^t : t \in \mathbb{N}_0\}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

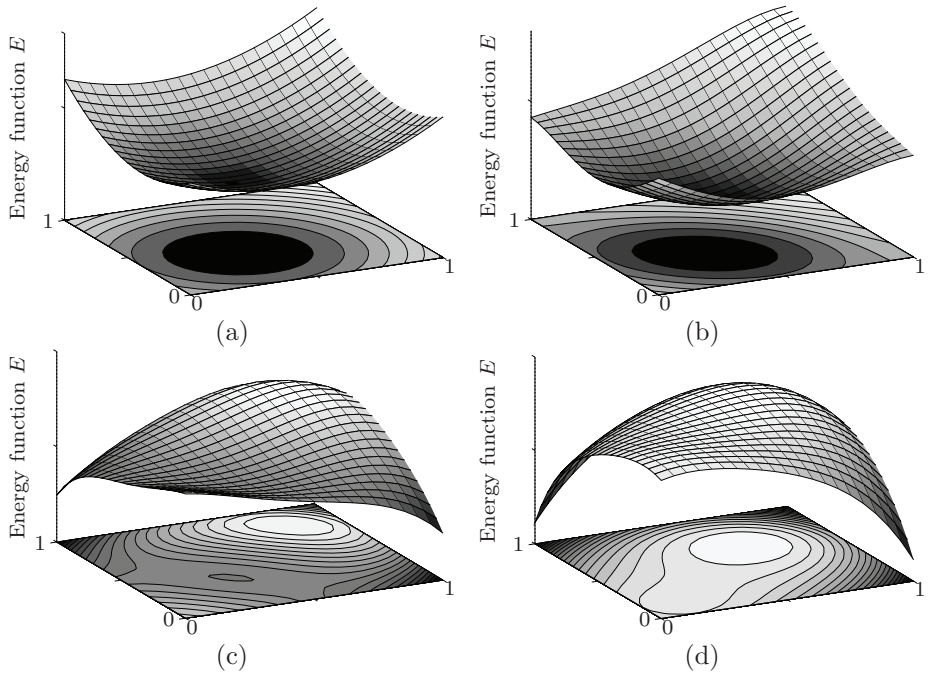

FIG. 3.2. *Four energy functions with two objects and increasing values of k. We have in the unit square* (a) *a unique minimum;* (b) *a unique minimum, but with other stationary points close to the boundary;* (c) *a unique minimum and other stationary points;* (d) *a unique maximum.*

Proof.

$$\begin{aligned}
[M^T - A^T \circ \mathbf{1}\boldsymbol{c}^T]^{\circ 2}\mathbf{1} &= \left(A^T \circ [M^T - \mathbf{1}\boldsymbol{c}^T]^{\circ 2}\right)\mathbf{1} \qquad \text{(by definition } A \circ A = A\text{)} \\
&= \left(A^T \circ [(M^T)^{\circ 2} - 2M^T \circ \mathbf{1}\boldsymbol{c}^T + (\mathbf{1}\boldsymbol{c}^T)^{\circ 2}]\right)\mathbf{1} \\
&= \left((M^T)^{\circ 2} - 2M^T \circ \mathbf{1}\boldsymbol{c}^T + A^T \circ (\mathbf{1}\boldsymbol{c}^T)^{\circ 2}\right)\mathbf{1} \\
&= [M^T]^{\circ 2}\mathbf{1} - 2M^T\boldsymbol{c} + A^T\boldsymbol{c}^{\circ 2}. \qquad \square
\end{aligned}$$

PROPOSITION 3.4. *If $k \in \mathcal{K}$, the system* (1.8), (1.9) *converges to the unique fixed point $\boldsymbol{r}^* \in \mathcal{H}$.*

Proof. First, we show that the energy function E decreases between any two iterations (see Figure 3.3); i.e., $E(\boldsymbol{r}^{t+1}) \le E(\boldsymbol{r}^t)$ for all $t \ge 0$. This is equivalent to proving that $(\boldsymbol{w}^{t+1})^T(\boldsymbol{w}^{t+1}) \ge (\boldsymbol{w}^t)^T(\boldsymbol{w}^t)$. Let us express $\boldsymbol{w}^{t+1}$ in terms of $\boldsymbol{w}^t$:

$$\begin{aligned}
\boldsymbol{w}^{t+1} &= \mathbf{1} - \frac{k}{n}[X^T - \mathbf{1}(\boldsymbol{r}^{t+1})^T]^{\circ 2}\mathbf{1} \\
&= \mathbf{1} - \frac{k}{n}[X^T - \mathbf{1}(\boldsymbol{r}^t)^T - \mathbf{1}(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^T]^{\circ 2}\mathbf{1} \\
&\quad \text{(by Lemma 3.3 with } A = \mathbf{1}\mathbf{1}^T,\ M^T = X^T - \mathbf{1}(\boldsymbol{r}^t)^T, \text{ and } \boldsymbol{c} = \boldsymbol{r}^{t+1} - \boldsymbol{r}^t\text{)} \\
&= \boldsymbol{w}^t + \frac{k}{n}\left(2\,(X^T - \mathbf{1}(\boldsymbol{r}^t)^T)(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t) - (\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^T(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)\mathbf{1}\right) \\
&= \boldsymbol{w}^t + \frac{k}{n}\,\boldsymbol{q},
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

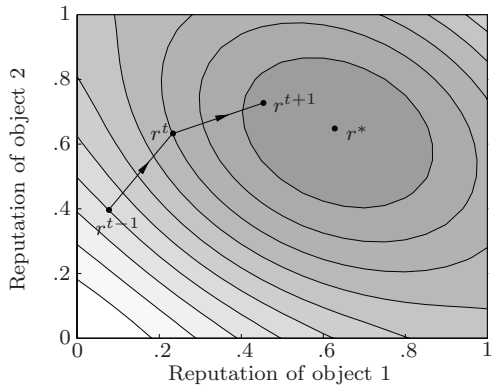

FIG. 3.3. *Two iteration steps by our method. Each decreases the energy function E. They take the steepest descent direction and converge to the minimum $\boldsymbol{r}^*$.*

with $\boldsymbol{q}:=\big(2\,(X^T-\mathbf{1}(\boldsymbol{r}^t)^T)(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)-(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)^T(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)\mathbf{1}\big)$. Hence,

$$\begin{aligned}(\boldsymbol{w}^{t+1})^T(\boldsymbol{w}^{t+1}) &= \left(\boldsymbol{w}^t+\frac{k}{n}\,\boldsymbol{q}\right)^T\left(\boldsymbol{w}^t+\frac{k}{n}\,\boldsymbol{q}\right)\\ &= (\boldsymbol{w}^t)^T(\boldsymbol{w}^t)+\frac{k^2}{n^2}\,\boldsymbol{q}^T\boldsymbol{q}+2\,\frac{k}{n}\boldsymbol{q}^T\boldsymbol{w}^t.\end{aligned}$$

Therefore, it is sufficient to show that $\boldsymbol{q}^T\boldsymbol{w}^t\geq 0$. This follows from

$$\begin{aligned}\boldsymbol{q}^T\boldsymbol{w}^t &= 2\,(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)^T(X-\boldsymbol{r}^t\mathbf{1}^T)\boldsymbol{w}^t-(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)^T(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)\mathbf{1}^T\boldsymbol{w}^t\\ &= 2\,(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)^T(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)\mathbf{1}^T\boldsymbol{w}^t-(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)^T(\boldsymbol{r}^{t+1}-\boldsymbol{r}^t)\mathbf{1}^T\boldsymbol{w}^t\\ &= \|\boldsymbol{r}^{t+1}-\boldsymbol{r}^t\|_2^2\,\mathbf{1}^T\boldsymbol{w}^t,\end{aligned}\tag{3.4}$$

which is greater than 0 since $\boldsymbol{w}^t$ is a positive vector because of the condition on k. This shows that the energy is strictly decreasing when $\boldsymbol{r}^{t+1}\neq\boldsymbol{r}^t$ with

$$E(\boldsymbol{r}^{t+1})-E(\boldsymbol{r}^t)\leq-\frac{\delta}{n}\,\|\boldsymbol{r}^{t+1}-\boldsymbol{r}^t\|_2^2,\tag{3.5}$$

where we use $\sum_{j=1}^m w_j>\delta$ for some $\delta>0$ (every weight is positive). Since E is lower bounded in $\mathcal{H}$, the sequence $(\boldsymbol{r}^t)$ converges to a single limit point $\boldsymbol{r}^*\in\mathcal{H}$. Then it follows from (1.13) that

$$\|\nabla_{\boldsymbol{r}}E(\boldsymbol{r}^t)\|_2=\frac{2\sum_{j=1}^m w_j}{n}\|\boldsymbol{r}^{t+1}-\boldsymbol{r}^t\|_2\leq\frac{2m}{n}\|\boldsymbol{r}^{t+1}-\boldsymbol{r}^t\|_2,$$

where we used $\sum_{j=1}^m w_j=m-k\sum_{j=1}^m d_j(\boldsymbol{r}^t)\leq m$. Therefore at $\boldsymbol{r}^*$, the gradient is zero, and by Proposition 3.2, the point $\boldsymbol{r}^*$ is the unique fixed point. ∎

3.3. The choice of parameter k. In Propositions 3.2 and 3.4 we assumed that the condition $k\in\mathcal{K}$ guarantees uniqueness or convergence. But, there exist larger values of k such that the minimum of E remains unique and the previous methods converge to this minimum. By increasing k such that $k\geq\sup\mathcal{K}$, we allow the maxima of E to appear in $\mathcal{H}$; see Figure 3.2(c). Then, we need to verify during the iteration steps whether $(\boldsymbol{r}^t)$ remains in the basin of attraction of the minimum $\boldsymbol{r}^*$ of E.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

PROPOSITION 3.5. *If the energy function E in* (3.1) *has a minimum, then the system* (1.8), (1.9) *is locally convergent and its rate of convergence is q-linear.*

Proof. We will prove that the function of iteration $F \circ G(\boldsymbol{r})$ is a contraction mapping in some neighborhood $\mathcal{N}$ of the minimum $\boldsymbol{r}^*$ of E. For that purpose, it is sufficient to show that $\|\nabla_{\boldsymbol{r}} F \circ G(\boldsymbol{r}^*)\|_2 < 1$. This gradient is given by

$$\nabla_{\boldsymbol{r}} \left(F \circ G(\boldsymbol{r}^*)\right) = \frac{2k}{n\, \mathbf{1}^T \boldsymbol{w}^*} (X - \boldsymbol{r}^* \mathbf{1}^T)(X - \boldsymbol{r}^* \mathbf{1}^T)^T, \tag{3.6}$$

where $\boldsymbol{w}^* = G(\boldsymbol{r}^*)$. Therefore, this gradient is positive semidefinite when $\mathbf{1}^T \boldsymbol{w}^* > 0$ (not guaranteed anymore since the weights can be negative). On the other hand, the Hessian of E at $\boldsymbol{r}^*$ must be positive definite (otherwise it contradicts the hypothesis of $\boldsymbol{r}^*$ as a minimum). The Hessian is given by

$$\begin{aligned} \nabla^2_{\boldsymbol{r}} E(\boldsymbol{r}^*) &= \frac{2\,\mathbf{1}^T \boldsymbol{w}^*}{n} \boldsymbol{I} - \frac{4k}{n^2} (X - \boldsymbol{r}^* \mathbf{1}^T)(X - \boldsymbol{r}^* \mathbf{1}^T)^T \\ &= \frac{2\,\mathbf{1}^T \boldsymbol{w}^*}{n} \left[\boldsymbol{I} - \nabla_{\boldsymbol{r}} \left(F \circ G(\boldsymbol{r}^*)\right)\right]. \end{aligned}$$

This implies that $\mathbf{1}^T \boldsymbol{w}^* > 0$ and the eigenvalues of $\frac{4k}{n^2}(X - \boldsymbol{r}^* \mathbf{1}^T)(X - \boldsymbol{r}^* \mathbf{1}^T)^T$ have to be strictly less than $\frac{2\,\mathbf{1}^T \boldsymbol{w}^*}{n}$. Therefore the spectrum of $\nabla_{\boldsymbol{r}} \left(F \circ G(\boldsymbol{r}^*)\right)$ belongs to $[0, 1[$ and hence its 2-norm is strictly less than 1.

In order to find the asymptotic rate of convergence, we consider the error at time k given by $\boldsymbol{\epsilon}^k := \boldsymbol{r}^k - \boldsymbol{r}^*$. This error is given using the first-order Taylor expansion at $\boldsymbol{r}^*$ by

$$\boldsymbol{\epsilon}^k = \nabla_{\boldsymbol{r}} \left(F \circ G(\boldsymbol{r}^*)\right)\, \boldsymbol{\epsilon}^{k-1} + O(\|\boldsymbol{\epsilon}^{k-1}\|^2).$$

Since the gradient in (3.6) is a symmetric matrix with all eigenvalues in $[0, 1[$, the asymptotic convergence is q-linear and the rate of convergence is $\rho\left(\nabla_{\boldsymbol{r}} \left(F \circ G(\boldsymbol{r}^*)\right)\right)$ for the Euclidean norm. ∎

The idea of increasing k is to make the discriminant function g more penalizing and therefore to have a better separation between honest and dishonest raters. A possibility is to take, during the first steps, the largest value of k such that the weights remain nonnegative; that is, at time t,

$$k^t : \min_j \boldsymbol{w}_j^t = 0, \tag{3.7}$$

where $\boldsymbol{w}^t$ is now given by

$$G(\boldsymbol{r}^t, k^t) = \mathbf{1} - k^t\, \frac{1}{n} \begin{pmatrix} \|\boldsymbol{x}_1 - \boldsymbol{r}^t\|_2^2 \\ \vdots \\ \|\boldsymbol{x}_m - \boldsymbol{r}^t\|_2^2 \end{pmatrix}.$$

In this manner, the worst rater at time t receives no weight for his own votes, and $\boldsymbol{w}^t$ is nonnegative. Usually, the sequence (k^t) converges, and one reaches the unique stable fixed point. This will be illustrated later by simulations.

4. Sparsity pattern and dynamical votes. This section extends some previous results to the case where the rating matrix has some sparsity pattern, that is, when an object is not evaluated by all raters. Moreover, we analyze dynamical rating matrices representing votes that evolve over time.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

4.1. Sparsity pattern. In general, the structure of real data is sparse. We hardly find a set of raters and objects with a vote for all possible pairs. An absence of a vote for object i from rater j will imply that the entry (i, j) of the matrix X is equal to zero; that is, by using the adjacency matrix A, if $A_{ij} = 0$, then $X_{ij} = 0$. These entries must not be considered as votes but instead as missing values. Therefore the previous equations presented in matrix form require some modifications that will include the adjacency matrix A. We write the new equations and their implications using the order of the previous section.

IF systems. The belief divergence for IF systems in (1.1) becomes

$$\boldsymbol{d} = \begin{pmatrix} \frac{1}{\boldsymbol{n}_1}\|\boldsymbol{x}_1 - \boldsymbol{a}_1 \circ \boldsymbol{r}\|_2^2 \\ \vdots \\ \frac{1}{\boldsymbol{n}_m}\|\boldsymbol{x}_m - \boldsymbol{a}_m \circ \boldsymbol{r}\|_2^2 \end{pmatrix}, \tag{4.1}$$

where $\boldsymbol{a}_j$ is the jth column of the adjacency matrix A and n_j is the jth entry of the vector $\boldsymbol{n}$ containing the numbers of votes given to each item, i.e., $\boldsymbol{n} = A^T\mathbf{1}$. The scalar n remains the total number of objects.

Quadratic IF systems. Equation (1.3) for quadratic IF systems will now be given in matrix form by

$$F(\boldsymbol{w}) = \frac{[X\boldsymbol{w}]}{[A\boldsymbol{w}]},$$

where $\frac{[\cdot]}{[\cdot]}$ is the componentwise division. Let us remark that every entry of $A\boldsymbol{w}$ must be strictly positive. This means that every object is evaluated by at least one rater with nonzero weight. With these modifications, the iterations and the fixed point are given by quadratic equations in $\boldsymbol{r}$ and $\boldsymbol{w}$:

$$\left(A \circ \boldsymbol{r}^{t+1}\mathbf{1}^T\right)\boldsymbol{w}^t = X\boldsymbol{w}^t, \tag{4.2}$$

$$\left(A \circ \boldsymbol{r}^{*}\mathbf{1}^T\right)\boldsymbol{w}^* = X\boldsymbol{w}^*. \tag{4.3}$$

Hence we expect an energy function to exist, and Theorem 1.3 is generalized by the following theorem.

THEOREM 4.1. *The fixed points of quadratic IF systems with integrable discriminant function g are the stationary points of the energy function*

$$E(\boldsymbol{r}) = \frac{1}{n}\sum_{j=1}^{m} n_j \int_0^{d_j(\boldsymbol{r})} g(u)\,du + c, \tag{4.4}$$

where d_j is the belief divergence of rater j that depends on $\boldsymbol{r}$ and $c \in \mathbb{R}$ is a constant. Moreover, one iteration step in quadratic IF systems corresponds to a dilated steepest descent direction with a particular step size

$$\boldsymbol{r}^{t+1} = \boldsymbol{r}^t - \boldsymbol{\alpha}^t \circ \nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t), \quad \text{with } \boldsymbol{\alpha}^t = \frac{n}{2}\frac{[\mathbf{1}]}{[A\boldsymbol{w}^t]}. \tag{4.5}$$

Proof. We have $\nabla_{\boldsymbol{r}} E(\boldsymbol{r}) = \nabla_{\boldsymbol{r}} \boldsymbol{d}^T \cdot \nabla_{\boldsymbol{d}} E(\boldsymbol{r})$ with

$$\nabla_{\boldsymbol{r}} \boldsymbol{d}^T = -2\,\frac{\left[X - A \circ \boldsymbol{r}\mathbf{1}^T\right]}{[\mathbf{1}\boldsymbol{n}^T]} \quad \text{and} \quad \nabla_{\boldsymbol{d}} E(\boldsymbol{r}) = \frac{1}{n}\left[\boldsymbol{n} \circ G(\boldsymbol{r})\right].$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Therefore a stationary point $\boldsymbol{r}^*$ in E satisfies

$$-\frac{2}{n}\,\frac{\left[X - A\circ \boldsymbol{r}^*\mathbf{1}^T\right]}{[\mathbf{1}\boldsymbol{n}^T]}\,(\boldsymbol{n}\circ G(\boldsymbol{r}^*)) = 0,$$
$$-\frac{2}{n}\,(X - A\circ \boldsymbol{r}^*\mathbf{1}^T)\,G(\boldsymbol{r}^*) = 0,$$
$$(A\circ \boldsymbol{r}^*\mathbf{1}^T)\,\boldsymbol{w}^* = X\boldsymbol{w}^*,$$

which corresponds to the fixed point equation given in (4.3).

We also have $\nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t) = \nabla_{\boldsymbol{r}}(\boldsymbol{d}(\boldsymbol{r}^t))^T \cdot \nabla_{\boldsymbol{d}} E(\boldsymbol{r}^t)$ with

$$\nabla_{\boldsymbol{r}}(\boldsymbol{d}(\boldsymbol{r}^t))^T = -2\,\frac{\left[X - A\circ \boldsymbol{r}^t\mathbf{1}^T\right]}{[\mathbf{1}\boldsymbol{n}^T]} \quad \text{and} \quad \nabla_{\boldsymbol{d}} E(\boldsymbol{r}^t) = \frac{1}{n}\left[\boldsymbol{n}\circ G(\boldsymbol{r}^t)\right] = \frac{1}{n}\left[\boldsymbol{n}\circ \boldsymbol{w}^t\right].$$

Therefore

$$\nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t) = -\frac{2}{n}(X - A\circ \boldsymbol{r}^t\mathbf{1}^T)\boldsymbol{w}^t \tag{4.6}$$
$$= -\frac{2}{n}(A\boldsymbol{w}^t)\circ(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t) = -\frac{\left[\boldsymbol{r}^{t+1} - \boldsymbol{r}^t\right]}{[\boldsymbol{\alpha}^t]}. \qquad \square$$

The number of votes n_j somehow gives a weight of importance for the minimization of the surface $\int_0^{d_j} g(u)\,du$. Therefore a rater with more votes receives more attention in the minimization process. Moreover, the table in section 2.3 becomes

Function g	$\boldsymbol{r}^*$ minimizes
$g(d) = \frac{1}{d}$	$\sum_j n_j \log d_j$
$g(d) = \frac{1}{\sqrt{d}}$	$\sum_j -n_j\sqrt{d_j}$
$g(d) = e^{-k\,d}$	$\sum_j -n_j e^{-k\,d_j}$
$g(d) = 1 - k\,d$	$\sum_j n_j d_j - n_j\frac{k}{2}\sum_j d_j^2$

Affine quadratic IF systems. The system for the discriminant function $g(d) = 1 - k\,d$ is given by

$$\boldsymbol{r}^{t+1} = F(\boldsymbol{w}^t) = \frac{[X\boldsymbol{w}]}{[A\boldsymbol{w}]}, \tag{4.7}$$

$$\boldsymbol{w}^{t+1} = G(\boldsymbol{r}^{t+1}) = \mathbf{1} - k\begin{pmatrix} \frac{1}{\boldsymbol{n}_1}\|\boldsymbol{x}_1 - \boldsymbol{a}_1\circ \boldsymbol{r}^{t+1}\|_2^2 \\ \vdots \\ \frac{1}{\boldsymbol{n}_m}\|\boldsymbol{x}_m - \boldsymbol{a}_m\circ \boldsymbol{r}^{t+1}\|_2^2 \end{pmatrix}, \tag{4.8}$$

starting with equal weights $\boldsymbol{w}^0 = \mathbf{1}$.

The energy function is given by

$$E(\boldsymbol{r}) = -\frac{1}{2\,kn}\,\boldsymbol{w}^T[\boldsymbol{w}\circ \boldsymbol{n}], \tag{4.9}$$

where $\boldsymbol{w}$ depends on $\boldsymbol{r}$ according to the function $G(\boldsymbol{r})$.

Proposition 3.2 remains valid for the system (4.7), (4.8), and the arguments are similar. The steepest descent method adapted to the system (4.7), (4.8) converges

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

with the property that the sequence $(E(\boldsymbol{r}^t))$ decreases. The proofs are closely related to those presented in Proposition 3.4.

PROPOSITION 4.2. *If $k \in \mathcal{K}$, the system* (4.7), (4.8) *converges to the unique fixed point $\boldsymbol{r}^* \in \mathcal{H}$.*

Proof. First, we show that the energy function E decreases between any two iterations; i.e., $E(\boldsymbol{r}^{t+1}) \leq E(\boldsymbol{r}^t)$ for all $t \geq 0$. By (4.9), this is equivalent to proving that $(\boldsymbol{w}^{t+1})^T(\boldsymbol{w}^{t+1} \circ \boldsymbol{n}) \geq (\boldsymbol{w}^t)^T(\boldsymbol{w}^t \circ \boldsymbol{n})$. Let us express $\boldsymbol{w}^{t+1}$ in terms of $\boldsymbol{w}^t$; then we obtain

$$\begin{aligned}
\boldsymbol{w}^{t+1} &= \mathbf{1} - \frac{[k\,\mathbf{1}]}{[\boldsymbol{n}]} \circ [X^T - A^T \circ \mathbf{1}(\boldsymbol{r}^{t+1})^T]^{\circ 2}\mathbf{1} \\
&= \mathbf{1} - \frac{[k\,\mathbf{1}]}{[\boldsymbol{n}]} \circ [X^T - A^T \circ \mathbf{1}(\boldsymbol{r}^t)^T - A^T \circ \mathbf{1}(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^T]^{\circ 2}\mathbf{1} \\
&\quad \text{(by Lemma 3.3 with } M^T = X^T - A^T \circ \mathbf{1}(\boldsymbol{r}^t)^T \text{ and } \boldsymbol{c} = \boldsymbol{r}^{t+1} - \boldsymbol{r}^t) \\
&= \boldsymbol{w}^t + \frac{[k\,\mathbf{1}]}{[\boldsymbol{n}]} \circ \left(2\,(X^T - A^T \circ \mathbf{1}(\boldsymbol{r}^t)^T)(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t) - A^T(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2}\right) \\
&= \boldsymbol{w}^t + \frac{[k\,\mathbf{1}]}{[\boldsymbol{n}]} \circ \boldsymbol{q},
\end{aligned}$$

where $\boldsymbol{q} := \left(2\,(X^T - A^T \circ \mathbf{1}(\boldsymbol{r}^t)^T)(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t) - A^T(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2}\right)$. Hence,

$$\begin{aligned}
(\boldsymbol{w}^{t+1})^T(\boldsymbol{w}^{t+1} \circ \boldsymbol{n}) &= \left(\boldsymbol{w}^t + \frac{[k\,\mathbf{1}]}{[\boldsymbol{n}]} \circ \boldsymbol{q}\right)^T (\boldsymbol{w}^t \circ \boldsymbol{n} + k\,\boldsymbol{q}) \\
&= (\boldsymbol{w}^t)^T(\boldsymbol{w}^t \circ \boldsymbol{n}) + \left(\frac{[k^2\mathbf{1}]}{[\boldsymbol{n}]} \circ \boldsymbol{q}\right)^T \boldsymbol{q} + 2k\,\boldsymbol{q}^T\boldsymbol{w}^t.
\end{aligned}$$

Therefore, it is sufficient to show that $\boldsymbol{q}^T\boldsymbol{w}^t \geq 0$:

$$\begin{aligned}
\boldsymbol{q}^T\boldsymbol{w}^t &= 2\,(\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^T(X - A \circ \boldsymbol{r}^t\mathbf{1}^T)\boldsymbol{w}^t - ((\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2})^T A\boldsymbol{w}^t \\
&= 2\,((\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2})^T A\boldsymbol{w}^t - ((\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2})^T A\boldsymbol{w}^t \\
&= ((\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)^{\circ 2})^T A\boldsymbol{w}^t.
\end{aligned}$$

Then, since every entry of $A\boldsymbol{w}^t$ is larger than some $\delta > 0$ by the condition on k, the energy is strictly decreasing when $\boldsymbol{r}^{t+1} \neq \boldsymbol{r}^t$:

$$E(\boldsymbol{r}^{t+1}) - E(\boldsymbol{r}^t) \leq -\frac{\delta}{n}\,\|\boldsymbol{r}^{t+1} - \boldsymbol{r}^t\|_2^2, \tag{4.10}$$

as, in addition, E is lower bounded in $\mathcal{H}$, and the sequence $(\boldsymbol{r}^t)$ converges to a single limit point $\boldsymbol{r}^* \in \mathcal{H}$. Then the gradient of E must be zero in $\boldsymbol{r}^*$ because we have by (1.13)

$$\begin{aligned}
\|\nabla_{\boldsymbol{r}} E(\boldsymbol{r}^t)\|_2 &= \frac{2}{n}\left\|A\boldsymbol{w}^t \circ (\boldsymbol{r}^{t+1} - \boldsymbol{r}^t)\right\|_2 \\
&\leq \frac{2}{n}\|\boldsymbol{r}^{t+1} - \boldsymbol{r}^t\|_2,
\end{aligned}$$

where we used that $A\boldsymbol{w}^t$ is componentwise upper bounded by $\mathbf{1}$. Therefore the gradient is zero at $\boldsymbol{r}^*$, and by Proposition 3.2 (also valid for our method with sparse matrix), the point $\boldsymbol{r}^*$ is the unique fixed point. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The choice of k can be made larger to better separate honest from dishonest raters, as discussed in section 3.3. Proposition 3.5 remains valid with a few modifications in its proof to take into account the adjacency matrix A.

This section has demonstrated that the earlier analysis can still be applied when we introduce a sparsity pattern in the rating matrix.

4.2. Dynamical votes. We consider in this section the case of time-varying votes. Formally, we consider discrete sequences

$$(X^s)_{s\geq 0}, \qquad (A^s)_{s\geq 0}$$

of rating matrices and adjacency matrices evolving over discrete time s. Let us remark that the scale time for that sequence must be chosen according to the frequency of new votes, e.g., every time we have some percentage of new votes. Hence the IF system (1.6), (1.7) takes into account the new rating matrix X^{s+1} in the functions F_{s+1} and G_{s+1} that become time-dependent:

$$\boldsymbol{r}^{t+1} = F_{s+1}(\boldsymbol{w}^t), \tag{4.11}$$

$$\boldsymbol{w}^{t+1} = G_{s+1}(\boldsymbol{r}^{t+1}). \tag{4.12}$$

The system (4.7), (4.8) for dynamical rating matrices is then given by

$$\boldsymbol{r}^{t+1} = F_{s+1}(\boldsymbol{w}^t) = \frac{\left[X^{s+1}\boldsymbol{w}^t\right]}{\left[A^{s+1}\boldsymbol{w}^t\right]}, \tag{4.13}$$

$$\boldsymbol{w}^{t+1} = G_{s+1}(\boldsymbol{r}^{t+1}) = \mathbf{1} - k \begin{pmatrix} \frac{1}{\boldsymbol{n}_1^{s+1}} \|\boldsymbol{x}_1^{s+1} - \boldsymbol{a}_1^{s+1} \circ \boldsymbol{r}^{t+1}\|_2^2 \\ \vdots \\ \frac{1}{\boldsymbol{n}_m^{s+1}} \|\boldsymbol{x}_m^{s+1} - \boldsymbol{a}_m^{s+1} \circ \boldsymbol{r}^{t+1}\|_2^2 \end{pmatrix}, \tag{4.14}$$

starting with equal weights $\boldsymbol{w}^0 = \mathbf{1}$.

We already know that for subsequent constant matrices X^s with $T_1 \leq s \leq T_2$, the iterations on $\boldsymbol{r}^t$ and $\boldsymbol{w}^t$ of system (4.13), (4.14) tend to fixed vectors $\boldsymbol{r}^*$ and $\boldsymbol{w}^*$, provided that k is not too large.

Each iteration on $\boldsymbol{r}^t$ decreases an energy function E^t dependent on the iteration step t. Intuitively, if the votes hardly change over time, then the sequence $(\boldsymbol{r}^{*t})$, where $\boldsymbol{r}^{*t}$ is the minimum of E^t, evolves slowly. Therefore the mapping will be mostly contracting: different initial conditions will eventually converge to the same trajectory that follows this sequence of minima. However, if we have significant changes in the votes, reflecting a lack of consensus in the opinions, then no such result of convergence can be established.

However, when the rating matrix is periodic, all the simulations we made always show convergent orbits for the trajectories $(\boldsymbol{r}^t)$; see, for example, Figure 4.1. Without going into technical details, we mention that it is still possible to prove the convergence for the specific case of 2-periodic sequences of rating matrices where the number of votes remains the same for every rater. That result comes from the existence of a polynomial of degree 4 as energy function similar to the one introduced before. Unfortunately, that function does not exist anymore when we increase the periodicity ($p > 2$) or when the number of votes differs from one step to another (see [17] for further details).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

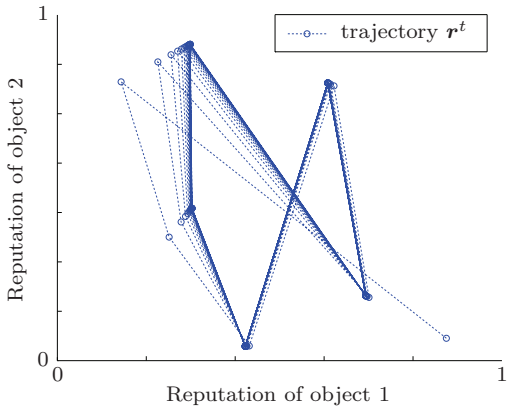

FIG. 4.1. *Trajectory of reputations (the circles) given by system* (4.13), (4.14) *with a periodic rating matrix* ($p = 5$) *and the evolving parameter* k^t *defined in* (3.7).

5. Simulations. We illustrate the general affine IF method in system (4.7), (4.8) with the updates on the parameter k given in (3.7). Two sets of data are used for this purpose:

1. the votes of 43 countries during the final of EuroVision 2008;
2. the votes of 943 movie lovers at the MovieLens web site.

We will see how and whom our method penalizes through the iterations. In the first set of data, we compare the difference in the ranking used by EuroVision and the ranking obtained by our method. The second set of data is used in order to verify the desired property mentioned in the introduction: *raters whose opinions often differ from those of other raters are given less weight.* For this purpose, we added fictional raters that diverge from the original raters.

5.1. EuroVision. During the final of EuroVision in 2008, 43 countries evaluated 25 songs. Each song represents a country, and no country among the 43 may vote for itself. Each country distributes its votes that are taken from $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 10, 12\}$, and only the vote zero can be given several times, while the others must be given exactly once.

We can expect that votes are driven by linguistic and cultural proximities between countries. However, vote alliances may appear between nearby countries by, for example, exchanging the maximal votes (as, e.g., discussed in [1]). Such raters should be penalized when they are, in addition, not in agreement with the other votes.

For this example we focus on the final ranking rather than on the reputations themselves. Clearly, the changes in the rankings due to different weights for the votes are sensitive and complex, but this final ranking is the main output of the competition and therefore its perturbation is crucial. We then compare in Figure 5.1 the ranking obtained by taking the average of the votes used by the jury of EuroVision and the ranking given by the affine IF method.

Results. Let us recall that the parameter k plays a role of discrimination between marginal and reasonable raters. For instance, small k's do not change the order in the original ranking, but increasing values of k make greater jumps appear in the ranking list. The parameter k is updated after every iteration step according to (3.7). After one iteration step, we have $k^1 = 0.06$, then $k^2 = 0.05$, and eventually it converges to

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

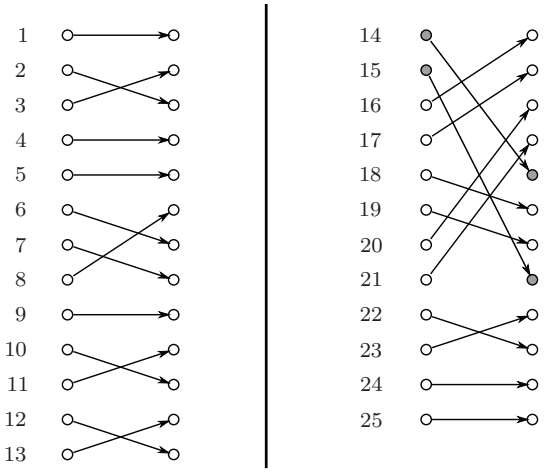

FIG. 5.1. *Difference of the rankings between the affine IF method and the average of the votes: for example, the second and third positions switch, and two countries (grey nodes) drop by four and six positions when we use the affine IF method.*

0.05. We briefly highlight two phenomena shown in Figure 5.1: the switch between the second and the third countries, and the loss of six positions for the 15th in the original ranking.

The switch occurs between Ukraine and Greece, which already had a small difference in the original ranking: averages of 5.5 and 5.2. These two countries received opposite votes from many raters. For instance, three times we find that one of those countries received 0 while the other received the maximum 12. This leads to significant changes when the votes of the raters are weighted. The switch shows that raters favorable to Greece were more objective according to our definition of the belief divergence.

Denmark dropped by six positions when we applied our method. Looking at the votes, the weights of supporters for this country are below the average. In particular, we point out that Denmark and Iceland gave 12 to each other. Then both received small weights for their votes; consequently the 12 they exchanged were assigned less weight in the second ranking (Iceland was 14th and dropped by four positions).

5.2. MovieLens. Our experiment concerns a data set[2] of 100,000 evaluations given by 943 raters on 1682 movies that range from 1 to 5. The data have been cleaned so that each rater rates at least 20 movies.

In order to test the robustness of our method, two types of behavior are analyzed in what follows: first, raters who give random evaluations, and second, spammers who try to improve the reputation of their preferred items.

Malicious raters. We added to the original data set 237 raters evaluating randomly a random set of items, or spammers always giving 1 except for their preferred movie, which they rated 5. In this manner, 20% of the raters are malicious. Let r^* and $\tilde{r}^*$ be, respectively, the reputation vector before and after the addition of the malicious raters. In this configuration, we expect that the malicious raters will receive smaller weights than the original raters and therefore the vectors of reputations should not be too different.

[2]The MovieLens data set used in this paper was supplied by the GroupLens Research Project.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

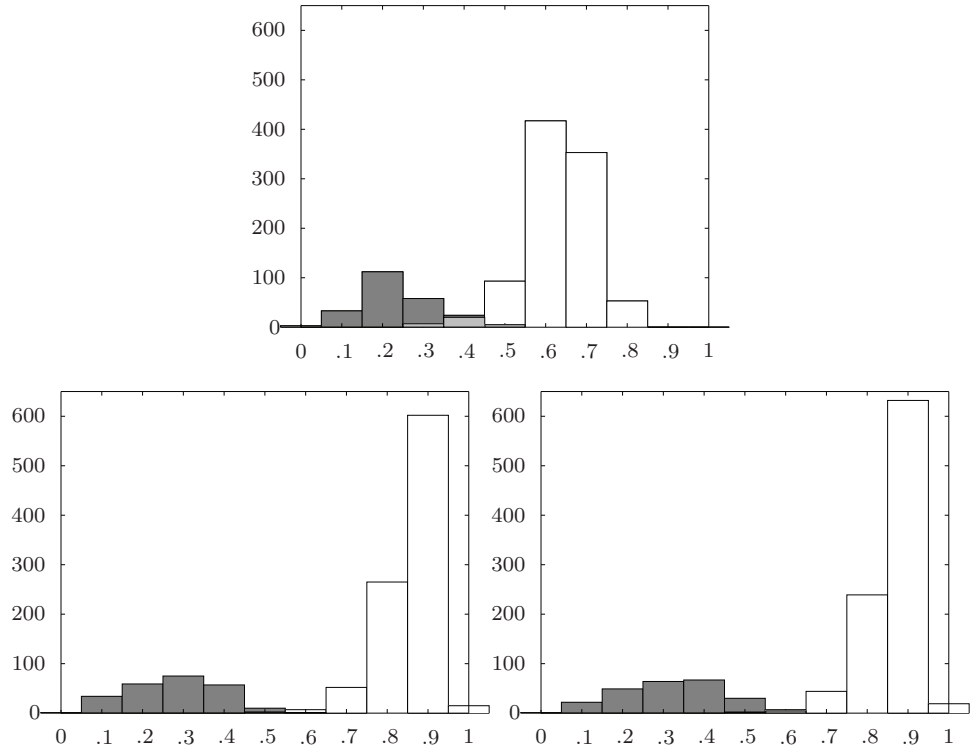

FIG. 5.2. *X-axis: the weights of the raters. Y-axis: the density after one iteration (top), after two iterations (left), and after convergence (right). In black: the spammers. In white: the original raters. In grey: overlap of both raters.*

We measure that difference by the distance $\|r^* - \tilde{r}^*\|_1 = 182$. This is naturally smaller than the same distance for the average method in which random raters receive as much weight as the others: $\|r^* - \tilde{r}^*\|_1 = 259$. For the case of spammers, we have a distance of 267 when using our method and a distance of 638 for the average method. These results are clearly a consequence of the weights given to the raters. Figure 5.2 shows the evolution of the weights during the iterations for the case of spammers. We remark that one iteration of the algorithm already separates spammers and original raters, but it is indeed useful to wait until convergence to have a better separation. Moreover, we see that a minority of raters in the original data have weights that are not better than those of random raters. Such raters diverge as much as random raters according to our method, and therefore they are penalized in a similar fashion. More figures can be found in [17, 18]. The parameter k is updated after every iteration step according to (3.7). After one iteration step, we have $k^1 = 0.03$, then $k^2 = 0.13$, and eventually it converges to 0.12.

6. Conclusion and further research.

6.1. Summary of the results. The general definition of IF systems provides a new framework for analyzing and evaluating voting systems. We emphasized the need for a differentiation of trusts between the raters unlike what is usually done on the Web. The originality of the approach lies in the continuous validation scale for the votes. Next, we assumed that the set of raters is characterized by various possible

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

behaviors including raters who are clumsy or partly dishonest. However, the outliers, being in obvious disagreement with the other votes, remain detectable by the system as shown in the simulations in the cases of alliances, random votes, and spammers.

Our paper focuses on a natural subclass of IF systems called quadratic IF systems, and we show the existence of an energy function that allows us to link a steepest descent to each step of the iteration. It then follows that the system minimizes the belief divergence according to some norm defined from the choice of the discriminant function. We also analyze several discriminant functions; some (the inverse and EM functions) were already introduced in the literature, while others (the exponential and affine functions) are new.

The main effort in this paper concerns the analysis of quadratic IF systems with the choice of the affine discriminant function. This choice is motivated by a statistical interpretation of the distribution of the votes and by the explicit condition on the parameter k that guarantees the uniqueness of the solution. This unique solution has the interesting property of maximizing the Euclidean norm of the weights of the raters. However, the convergent solution found via the IF method represents the "truth" better or is not known, and, therefore, the comparison with other methods is certainly one of the perspectives for this topic. Finally, the analysis of the system, and, more precisely, its convergence, becomes tractable. We also give experimental results on real data sets that illustrate the relevance of our approach.

6.2. Perspectives. The quadratic IF systems already give a large choice of different systems via the discriminant function. The three candidates we discussed can be extended to other possibilities motivated by some statistical assumptions on the votes, or by the minimization of some energy function (that would also maximize some function on the weights). Moreover, by accepting complex dynamics for evolving votes, we can include multiple solutions in the system. As already mentioned, such equilibria may be interpreted as coherent opinions among groups of people, and such methods can be compared with clustering methods. In addition, it remains useful to keep a parameter in the discriminant functions (k in our examples) that allows us to adjust the level of separation.

Our validation tests can be expanded to other data sets, in particular, we think, to dynamical data sets that have been hardly investigated for voting systems. Moreover, behaviors other than those we presented are possible. We give naive profiles to the fictitious raters we added in the second simulation; however, we know that intermediate behaviors are much harder to detect. Despite everything, malicious people could use the system to vote correctly on several items and then cheat on a few items. Such a strategy can be discouraged when votes are not free (for example, the votes in eBay are made after a transaction) or by tracking these kinds of *traders*. In addition to experiments on other data sets with other profiles of raters, the comparison between different IF systems provides another area to pursue. In our case, we give a limited example with three raters and two ice skaters to compare three discriminant functions. Then, for larger data sets, we limit ourselves to the affine function, the only one that gives a unique solution, and we compare it with the average when some perturbations occur. If we disregard the property of uniqueness for the solution, we can compare our method with others, but then we need to discuss the choice of different initial points that lead to different solutions for these systems. On the other hand, what we want to compare is not trivial since the goal of IF systems is not a simple separation between good and bad raters but rather a continuous validation scale. For that reason, the tractability of affine IF systems makes it attractive.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Acknowledgment. We thank the GroupLens Research Project for supplying the MovieLens data set.

REFERENCES

[1] V. Ginsburgh and A. Noury, *Cultural Voting: The EuroVision Song Contest*, mimeo, 2004; available online from http://ideas.repec.org/p/cor/louvco/2005006.html#download.

[2] L. Page, S. Brin, R. Motwani, and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web*, Technical report, Stanford InfoLab, 1999; available online from http://ilpubs.stanford.edu:8090/422/.

[3] L. Mui, M. Mohtashemi, and A. Halberstadt, *A computational model of trust and reputation*, in Proceedings of the 35th Annual Hawaii International Conference on System Sciences, 2002, pp. 2431–2439.

[4] S. Kamvar, M. Schlosser, and H. Garcia-molina, *The Eigentrust algorithm for reputation management in P2P networks*, in Proceedings of the 12th International Conference on World Wide Web, ACM, New York, 2003, pp. 640–651.

[5] M. Richardson, R. Agrawal, and P. Domingos, *Trust management for the semantic web*, in The SemanticWeb—ISWC 2003, Lecture Notes in Comput. Sci. 2870, Springer, Berlin, Heidelberg, 2003, pp. 351–368.

[6] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, *Propagation of trust and distrust*, in Proceedings of the 13th International Conference on World Wide Web, ACM, New York, 2004, pp. 403-412.

[7] J. O'Donovan and B. Smyth, *Trust in recommender systems*, in Proceedings of the 10th International Conference on Intelligent User Interfaces, ACM, New York, 2005, pp. 167–174.

[8] G. Theodorakopoulos and J. Baras, *On trust models and trust evaluation metrics for ad hoc networks*, IEEE J. Sel. Areas Commun., 24 (2006), pp. 318–328.

[9] S. Zhang, Y. Ouyang, J. Ford, and F. Makedon, *Analysis of a low-dimensional linear model under recommendation attacks*, in Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2006, pp. 517–524.

[10] E. Kotsovinos, P. Zerfos, N. M. Piratla, N. Cameron, and S. Agarwal, *Jiminy: A scalable incentive-based architecture for improving rating quality*, in Trust Management, Lecture Notes in Comput. Sci. 3986, Springer, Berlin, Heidelberg, 2006, pp. 221–235.

[11] Z. Gyöngyi and H. Garcia-Molina, *Link spam alliances*, in Proceedings of the 31st International Conference on Very Large Data Bases (VLDB '05), 2005, pp. 517–528.

[12] R. Baeza-Yates, C. Castillo, and V. López, *PageRank increase under different collusion topologies*, in Proceedings of the First International Workshop on Adversarial Information Retrieval on the Web, 2005; available online from http://airweb.cse.lehigh.edu/2005/baeza-yates.pdf.

[13] G. M. Del Corso, A. Gullí, and F. Romani, *Ranking a stream of news*, in Proceedings of the 14th International Conference on World Wide Web, ACM, New York, 2005, pp. 97–106.

[14] P. Laureti, L. Moret, Y.-C. Zhang, and Y.-K. Yu, *Information filtering via iterative refinement*, Europhys. Lett., 75 (2006), pp. 1006–1012.

[15] Y.-K. Yu, Y.-C. Zhang, P. Laureti, and L. Moret, *Decoding information from noisy, redundant, and intentionally distorted sources*, Phys. A, 371 (2006), pp. 732–744.

[16] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley & Sons, New York, 1996.

[17] C. de Kerchove, *Ranking Large Networks: Leadership, Optimization, and Distrust*, Ph.D. thesis, Université Catholique de Louvain, Louvain-la-Neuve, Belgium, 2009.

[18] C. de Kerchove and P. Van Dooren, *Reputation systems and optimization*, SIAM News, 41 (2) (2008), pp. 1, 3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.