Data Mining and Knowledge Discovery (2021) 35:2251–2281
https://doi.org/10.1007/s10618-021-00783-3

Social media as author-audience games

Andre F. Ribeiro[1]

Received: 5 October 2018 / Accepted: 21 July 2021 / Published online: 10 August 2021

Abstract
We present an approach for the prediction of user authorship and feedback behavior with shared content. We consider that users use models of other users and their feedback to choose what to publish next. We look at the problem as a game between authors and audiences and relate it to current content-based user modeling solutions with no prior strategic models. As applications, we consider the large-scale authorship of Wikipedia pages, movies and food recipes. We demonstrate analytic properties, authorship and feedback prediction results, and an overall framework to study content authorship regularities in social media.

Keywords Content models · User models · Authorship · Game theory · Social media · Distance-metric learning · Text understanding

1 Introduction

We study new tools to model user authoring behavior in online media. Developing tools to predict and understand how users publish content in the presence of others and previous content is of practical relevance in both the design of information sharing systems and their attached applications, such as recommender systems.

Many regularities have been found in the online feedback behavior of users (Goyal et al. 2010; Radinsky et al. 2012; Szabó and Huberman 2010; Lerman 2007; Das and Lavoie 2014), but much fewer patterns have been discovered in online content creation. By remaining agnostic to the surrounding media (its author base constitution, incentives, audience, practices, etc.), standard topic and content models (Pennacchiotti and Gurumurthy 2011; Hu et al. 2015; Hong and Davison 2010; Cha and Cho 2012) might miss patterns relevant to understanding and predicting user behavior. There is a well-known strategic model for information propagation in social networks (the

Responsible editor: Aristides Gionis.

✉ Andre F. Ribeiro
andre_ribeiro@hks.harvard.edu

[1] University of Cambridge, Cambridge, USA

Table 1 Data and document representation used in the three studied scenarios

Corpus	Documents	Authors	Audience Feedback	Author Feedback	Document Representation (bag of)
Wikipedia	30M pages	Wikipedians	16M page views	100K-1M edit reversals	actions
Mainstream Movies	4.1K scripts	Directors	dollar revenue	–	actions
Yummly	369K recipes	Yummly Users	1.1M yums	–	ingredients

Behavior Contagion Model Degroot 1974; Nisan et al. 2007, section 2) that describes user feedback behavior formally. There is, however, no correspondent model for content creation. Content creation and authorship are computer-supported cooperative or competitive activities undertaken by millions everyday, but they have been rarely looked strategically as such. Our current goal is to demonstrate that current state-of-the-art techniques for content-based behavior prediction could benefit from prior models encompassing the strategic dimensions of online authorship.

Consider Wikipedia as an example. The problem in this case is to predict which and how many pages individual users are likely to edit in the future. Mathematically, we frame the problem as a decision problem for users and consider how expected feedback from others can constraint user authoring decisions. The relationship between authorship and predicted (positive or negative) feedback is an obvious aspect of everyday communication (e.g., this article, a book, a joke) but is especially apparent in social media. In Wikipedia, feedback consists mostly of corrections from others. Consider that users like publishing but not being corrected. As consequence, when few users are networked, their likelihood of being corrected is lower, so they have incentives to author topically widespread, but possibly erratic, pages and edits. In larger numbers, edits are falsified more and more aggressively, and users are required to be more and more precise and specialize. If we look at Wikipedia this way, its byproduct is a medium that increases in precision with an increasing user base.

We consider the three example corpora in Table 1, corresponding to texts and user bases of different sizes. We formulate an equilibrium between authors and readers, which is useful to predict the behavior of either side (authors or audiences). We then employ the proposed game to solve the following problems: predict how much and over which topics users will publish (**user authorship prediction**) and feedback (**user feedback prediction**) in the future. Input data consist of a set of texts, author ids and feedback counts (from audiences or other authors). Despite the formal model, the main question addressed is therefore practical: whether is possible to exploit regularities such as the previous (formally described by a game) to predict user behavior.

1.1 Model summary

Before discussing this problem and related work in detail, we outline the proposed model with a two-player only example. Let $\mathcal{D}^v$ be a v-dimensional metric space among documents and $y \in \mathcal{D}^v$ an individual document. We consider $S_i(y)$, the likelihood of user i authoring or posting y. We ask how posts from a second user j change i's posting behavior. Let a parameter β take values in $[-1, +1]$, with -1 for user content 'contagion', +1 'dispersion', and 0 no mutual relationship. Let then user i's set of k past documents be $U_i = \{y_0^i, y_1^i, \ldots, y_k^i\}$ which is associated with a mean position x_i and Covariance Matrix M_i (resp. j, U_j, M_j).

The game formulated is a formal argument, and a many-players generalization, for the following Gaussian form for $S_i(y)$ over $\mathcal{D}^v$:

$$\begin{aligned} S_i(y; \beta) &= \mathcal{N}(x_i, (M_i^{-1} - \beta M_j^{-1})), \\ &= \mathcal{N}(x_i, \Sigma_i), \end{aligned} \tag{1}$$

where $x_i \in \mathcal{D}^v$ is the mean of positions U_i and $\Sigma_i = M_i^{-1} - \beta M_j^{-1}$. The position x_i and matrix Σ_i thus expresses players' topical centrality and authorship 'domain' in the presence of others. When $\beta = 0$, players' distributions are uncorrelated. The distribution becomes the no-prior Maximum Likelihood Gaussian estimate over $\mathcal{D}^v$. The proposed model is therefore associated with the rejection of the hypothesis $\beta = 0$. When $\beta > 0$, players are more likely to post in areas distant from other players' posts (that are 'unattended') while still close to their own. The formulation transforms the user authorship prediction problem into a probability density estimation one, with an across-author parameter β. The relationship in Eq. 1 has also a straightforward interpretation in terms of a Mahalanobis distance metric (Bishop 2006; Goldberger et al. 2004; Kostinger et al. 2012) among players, x, and content, y. Learning a metric in such approaches is often formulated as a combination of Covariance Matrices assembled from data. In the present interpretation, players have their own Covariance Matrices, and therefore individual metrics, but share β. The parameter β is a 'macro-variable' that serves to, at the same time, describe and predict behavior in the medium from the easy-to-assemble player Covariance Matrices.

We also formulate a Hierarchical Bayesian Model to estimate β and Covariance Matrices Σ_i simultaneously from training datasets with multiple players (assigning them Uniform and Wishart priors respectively). The model thus also mitigates the assumption of noiseless Covariance Matrices in common Metric Learning approaches (Goldberger et al. 2004; Kostinger et al. 2012; Weinberger and Saul 2009).

The proposed framework has therefore two closely related parts: a new Game-theoretic model and a Bayesian generative model for content. The first is a formal model for the social behavior of authors, denoted the Predator-Watch Model (PWM), the second formulates how components of the first model can be estimated from real-world authorship or feedback observations. Using metric distances in the game and casting the estimation problem as a metric-learning problem ease the exposition in both fronts.

After estimating the previous parameters across problems, we consider how accurately $S_i(y; \beta)$ predicts users' authoring behavior in held-out data. We consider several text-based tasks: predict authors given a Wikipedia page, predict Wikipedia authors behavior (will a user edit a page?), predict movie audiences' feedback behavior and predict feedback behavior in a large recipe sharing site. We compare performance to LDA-based authorship models and other past solutions.

1.2 Reader roadmap

We start by discussing related work. In particular, we review how social behavior models and content modeling are typically connected (Sect. 2 *Background and Related Work*). We then turn to a new game-theoretical model of behavior in social media (Sect. 3 *The Predator-Watch Model (PWM)*). This is a mostly theoretical section. It will motivate and justify the ensuing content modeling approach, including the model and parameters that are later estimated. Users interested exclusively in the proposed techniques may skip this section at first. The methodological sections that follow then present the recommended Document Representation (Sect. 4, *Game and Document Representation)*, Document Metric (Sect. 5 *Metric Learning*), Model parameter estimation (Sect. 5.2 *Parameter Estimation*) and implementation details (Sect. 6 *Complexity and Implementation*). Finally, we use the proposed techniques to study authorship and content exchange in Wikipedia, the Movie industry and a Recipe sharing website (Sect. 7 *Experiments*).

2 Background and related work

A single user's decision to post in a medium depends simultaneously on the user's relationship to the medium's current content and the user's model of other users. Models for these two problems, **User-Content** and **User-User models** (Fig. 1a), have been developed across distinct research areas - with the combination of the two in particular receiving little attention. A popular statistical **User-Content Model** is the Latent Dirichlet Allocation (LDA). Let $P_i(w)$ be the probability of user i uttering a word w. A natural first step is to consider the probability that the user will utter w, $P(w|i)$, given he or she decides to communicate, $P(i)$. The latter describes a user's prior likelihood to communicate. LDA-like generative approaches attempt to estimate the former distribution by assuming a latent set Z of user topics,

$$\begin{aligned} P_i(w) &= P(i)P(w|i) \\ &= P(i)\sum_{z\in Z} P(w|z)P(z|i), \end{aligned} \tag{2}$$

and estimating, in turn, $P(w|z)$ and $P(z|i)$ from past or training text posts. LDA therefore decomposes user distributions, $P_i(w)$, into **topic-word**, $P(w|z)$, and **user-topic**, $P(z|i)$, distributions.

User interests, authoring/feedback prediction, etc. then follow from $P_i(w)$. Authorship prediction, in particular, has been addressed in the Author-Topic Model (ATM), Dirichlet-Multinomial Regression (DMR) and other approaches by adding document correlates or features to LDA (Pennacchiotti and Gurumurthy 2011; Hong and Davison 2010; Cha and Cho 2012; Rosen-Zvi et al. 2004).

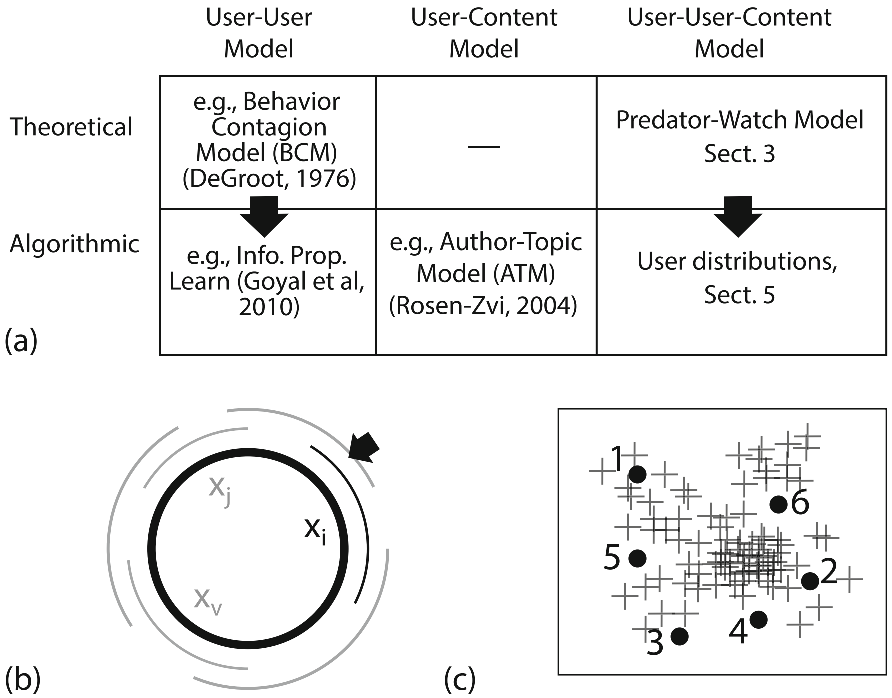
Social network models of feedback (Szabó and Huberman 2010; Lerman 2007) and content-propagation (Goyal et al. 2010; Radinsky et al. 2012; Franks et al. 2014), **User-User models**, stipulate patterns (of interdependence) among users' behavior but often ignore the semantic content of exchanged messages. Some have shown surprising accuracy and parsimony when predicting user feedback from those of neighbors, without modeling content. Beyond the algorithmic level, most of these follow, often implicitly, the Behavior Contagion (BCM) (Degroot 1974; Nisan et al. 2007) view of networks. The BCM has been studied extensively both in Economics (Bala and Goyal 1998; Degroot 1974) and in the Multi-Agents literature (Korkmaz et al. 2014; Olfati-Saber et al. 2007; Bosse et al. 2013; Grandi et al. 2015). The BCM is the central game-theoretic model of behavior in social networks. There is also a history of games in the study of both networks and language (Altman et al. 2006; Nisan et al. 2007; Borgs et al. 2011; Benz et al. 2006; Wang and Gasser 2002; Steels 2012). As the best framework we know to model inter-personal interaction, game theory should become increasingly important in the study of technology-mediated content creation as well.

The BCM stipulates that behavior contagion, a type of social influence, is the main driver of behavior in networks. Contagion is the propensity for certain behavior exhibited by one person to be copied by his or her neighbors. The BCM is especially well-suited to describe feedback behavior of networked individuals (Szabó and Huberman 2010; Lerman 2007; Korkmaz et al. 2014), which is often characterized by viral feedback cascades. In the game, a player i starts with a noisy signal, $P_i = \mathcal{N}(x_i, \Sigma_i)$, where $\Sigma_i \in \mathbb{R}^n$ is an error term whose components have zero mean and Normal distributions. This signal is therefore multi-dimensional on players and often unidimensional or binary in 'content' (e.g., how much players like some content or an opinion spectrum). The resulting game is then described by n individual distributions, $\mathbf{P}^t(w) = (P_1^t(w), ..., P_n^t(w))$, at time t. In the simplest case, player distributions are updated to increasingly resemble those of neighbors,

$$P_i^1(w) = \sum_{j=0}^{n} q_{ij} P_j^0(w), \tag{3}$$

where q_{ij} is a function (often stochastic and Bayesian-based) expressing j's influence in i. Sometimes the function is based on individuals' (self and/or others) precision (Σ_i^{-1}) (Demarzo et al. 2003). The model has both theoretic and practical ramifications (Bala and Goyal 1998; Goyal et al. 2010), Fig. 1a, and has been used to study both the short-term behavior of networked players, and the long-term equilibrium of player distributions, under several conditions (network connectivity (Olfati-Saber et al. 2007; Demarzo et al. 2003), learning procedures (Bala and Goyal 1998; Panait and Luke 2005; Chamley 2004), trust (Tsang and Larson 2014; Grandi et al. 2015), etc.)

Noticeably, however, the previous two models make complementary generative assumptions about user content. LDA assumes that user content is generated from a set of shared and stationary topics. Meanwhile, the BCM assumes that user opinion, belief or interest (and thus ultimately content) is generated from the opinions, beliefs or interests of neighbors. Notice that content plays no role in the BCM (only social influence and connectivity). The BCM is, after all, a model of content propagation (or 'audiences' in the framework below), with content creators often operating under other (content-based) incentives like novelty and specialization.



	User-User Model	User-Content Model	User-User-Content Model
Theoretical	e.g., Behavior Contagion Model (BCM) (DeGroot, 1976)	—	Predator-Watch Model Sect. 3
Algorithmic	e.g., Info. Prop. Learn (Goyal et al, 2010)	e.g., Author-Topic Model (ATM) (Rosen-Zvi, 2004)	User distributions, Sect. 5

Fig. 1 **a** Models of user content and mutual social influence have been studied extensively both in theory and applications, we consider a model where users model other users' content, **b** illustration of the main elements of the proposed authorship game (Predator Watch Model, PWM): author positions x_i over an unknown topical space $\mathcal{D}^1$ and audience viewership events (arrow), **c** Past page (+) and user (●) positions in the Wikipedia 'Living People' Category, 01/01/2014, Wikipedia authors often self-disperse topically, developing mutually complementary areas of expertise and interest

We thus consider whether there are advantages to studying models of content and behavior under a common framework for user authorship, **User-User-Content models**. We assume users shape $P_i(w)$ not from stationary topics or interests, but the topics or interests of others sharing the medium (be they audience members or other authors). We approach the estimation of $P_i(w)$ by stipulating a prior behavior model like the BCM, but one that (1) encompasses content, Fig. 1a, (2) consider authorship as a strategic game between authors and audiences, and (3) does not assume strict contagion - having a parameter β that indicates contagion to dispersion. This will couple user distributions with a prior and parametric strategic model for authorship, leading to closed-form distributions, $P_i(w)$, that are easy to train with authoring and feedback data.

To that end, we assume that topics are sampled from, instead of a discrete set of topics Z, a v-dimensional metric space $\mathcal{D}^v$,

$$\begin{aligned} P_i(w) &= P(i) \int_{y_k \in \mathcal{D}^v} P(w|y_k)P(y_k|i), \\ P(y_k|i) &\sim \mathcal{N}(x_i, \Sigma_i), \quad (user - topic) \end{aligned} \tag{4}$$

where $x_i \in \mathcal{D}^v$ is the (mean) position of user i's past documents and Σ_i his topic variance.[1] Namely, this assumes a metric among documents $d(y_k, y_v)$, where $y_k, y_v \in \mathcal{D}^v$. And this will suggest that users model not only other individuals, like in the BCM, but their high-dimensional positions in $\mathcal{D}^v$ (and thus, their mutually shared content). The assumption that $\mathcal{D}^v$ is a metric space follows the Metric Learning literature (Bellet et al. 2013; Bishop 2006). This will define a user-specific Mahalanobis metric based on his or hers distribution $\mathcal{N}(x_i, \Sigma_i)$.

3 The predator-watch model (PWM)

3.1 Game summary and background

We formulate a new game for online authorship where authors use (topical) models of others to choose content. We frame the authorship problem as follows. Individuals can either be authors (content-creators) or audience members (content-viewers). Author publications or posts are subject to two sources of feedback: viewership from the shared audience (positive) and corrections from other authors (negative).[2] Considering space $\mathcal{D}^v$, and that the position of users are positions of their posted documents in $\mathcal{D}^v$, the problem becomes: where in this space will authors publish given they can observe the positions of other authors and of audience members? Authors can, for example, choose to cover many topics superficially or focus on a few in detail. We formulate a solution to this problem where authors maximize viewership while minimizing corrections (author decisions thus depending on who they are sharing the medium with). This multi-party decision problem can be solved with game theory. A new game is first formulated in the interest of formal precision and communication. But it will also directly motivate the authorship prediction approach in Sect. 5.

The game is at a description level similar to the BCM, making explicit user incentives and their consequent collective behavior. Game solutions and representations are two fundamental concepts in game theory. Game solutions often take the form of an equilibrium. Like in other equilibria (e.g., chemical), a game equilibrium corresponds to a state of no-change in the game. Under a Nash equilibrium, in particular, a player does not gain anything from deviating from her chosen strategy, assuming the other players keep their strategies unchanged. The equilibrium therefore formulates player decisions when they are making them at the same time and the decision

[1] Notice that we use x for user and y for documents but they are both positions in $x, y \in \mathcal{D}^v$.

[2] Notice we use the term 'correction' as an abstraction for the 'amount' of observable or unobservable competition among authors.

of one player takes into account the decisions of other players. Potential games are games where player payoffs (a.k.a., utilities or gains) can be described by a continuous potential function (Milchtaich 1996; Nisan et al. 2007)(Sect.19), a distance metric being an example. The name and intuition behind these games come from typical physical potential functions (e.g., gravity), where a potential function between two points depend solely on their absolute distance. In the game setting, this implies that distances in some prior space are sufficient to describe players' utilities. The previous player-content metric assumptions therefore suggest this subclass of games. Potential games have specific and well-known equilibria properties, and we describe a solution in this framework.

Games with uncertainties are often represented with a probabilistic simplex. Let A be the set of actions players can take, which is often called the game pure strategy set. Define $\mathcal{D}_A^v$ as the simplex of A, the set of discrete probability distributions over A,

$$\mathcal{D}_A^v : \{y \in \mathbb{R}^v : y(0) + \cdots + y(v-1) = 1, y(a) \geq 0, a = 0, \ldots, v-1\}. \tag{5}$$

A mixed strategy for player i is, in turn, a position in this space, $x_i \in \mathcal{D}_A^v$, whose components define the probability that the player will take each action $a \in A$ (e.g., utter the word a). We next formulate the game and its solution. Later, we will define a simplex-based representation for documents, which serves as game representation and topic-word distributions. For a more extensive overview of social behavior and game-theory, we recommend (Nisan et al. 2007) (in particular, chapters 19 and 24).

3.2 The predator-watch model

Consider the following abstract game. A tribe is vulnerable to predator attacks and detecting them reliably is of great value. The game consists of n players, who can communicate with alarm calls, one predator that can attack from any position (with an unknown distribution), and is played over a v-dimensional metric space $\mathcal{D}_A^v$. In this abstract game, authors correspond to tribe members and the audience to the predator whose time-changing position tribe members jointly try to predict. We reserve the word 'player' to refer to authors (defining the 'predator' as a single player that encapsulates the entire audience).

A player i can position himself around the tribe perimeter and survey for the predator. His strategic choice $\mathcal{N}(x_i, \Sigma_i)$ consists then of a position, $x_i \in \mathcal{D}_A^v$, and the area he surveys, a dispersion matrix $\Sigma_i \in \mathbb{R}^{v \times v}$. Table 2 summarizes the notation used and Fig. 1b illustrates informally the game main elements (with player areas shown as circle segments and a single attack as an arrow). A player can cover the entire perimeter (all directions) but, as consequence, will be erratic. Players can, instead, cover smaller areas and inter-communicate. With an increasing player count n, players can detect the predator this way with increasing precision. Like the BCM, the game is therefore described by player distributions, $\mathbf{S}^t(y) = \{S_1^t(y), \ldots, S_n^t(y)\}$.

We study this game as an abstraction for social-media authorship at a given point in time. We take predator spottings as audience views [3] and player-covered areas as

[3] An event where an audience member views some content, such as Wikipedia page views.

Table 2 Model notation summary

Symbol	$\in$	Description
$\mathcal{D}^v_A$	$\mathbb{R}^v$	v-dimensional simplex the game is played over, its dimensions A are an abstract set of words that capture shared content (see Sect. 4)
y_k	$\mathcal{D}^v_A$	mixture of words and document k's position in $\mathcal{D}^v_A$
x_i	$\mathcal{D}^v_A$	player i's position in $\mathcal{D}^v_A$, the mean position of his or her past documents.
$S_0(y)$	$[0, +1]$	audience's probability of reading y
$S_i(y)$	$[0, +1]$	author i's probability of authoring y, $0 < i \leq n$
Σ_i	$\mathbb{R}^{v \times v}$	player i's topic range
$\vert\Sigma_i\vert$	$\mathbb{R}^+$	player i's area, the norm of Σ_i and a semantically-informed estimate for i's topic range.
β	$[-1, +1]$	across-author incentive for topic specialization.
$\lambda^t(i, j)$	$\mathbb{R}^+$	payoff for player i in a 2-player game with j at time t.
$w_1(y)$	$\{1, 2, \ldots, n\}$	the closest player to y.
$w_2(y)$	$\{1, 2, \ldots, n\}$	the second-closest player to y.

players' topics. Figure 1c shows past pages ('spottings') as crosses in the 'Living people' Wikipedia Category and the positions of its 6 most active authors as circles (01/01/2014, with the learned metric, projected to 2D with PCA). The abstraction is interesting because it articulates, with minimal elements, that to choose a position x_i, a player must not only take into account some externality (the predator's position), but also all other players (the positions of all networked individuals). We start by framing the game as a statistical problem, followed by its utility structure, equilibria and parameters.

3.3 Area, precision and payoffs

The player's chosen position, x_i, serves as a hypothesis (with an alarm bringing forward the hypothesis that the predator is at x_i). We are interested in the random variable $Y \in \mathcal{D}^v_A$ of predator attack positions. We use $y \in \mathcal{D}^v_A$ for both a generic document position or an audience view (predator) position, according to context. Suppose then there are two players, i and j, and thus two alternative hypotheses about the variable's probability distribution. Let's assume that the predator plays with density $S_0(Y)$, and players i and j with $S_i = \mathcal{N}(x_i, \Sigma_i)$ and $S_j = \mathcal{N}(x_j, \Sigma_j)$. Notice that we index the predator with 0 and players $0 < i, j \leq n$.

We relate the game's three main elements (player area, precision and payoffs) with these distributions. The relationship between the first two corresponds to the familiar relationship between players' Covariance and Precision Matrices. If a player chooses a wide area $\|\Sigma_i\|$ (which may allow him to detect many attacks), he will suffer in precision $\|\Sigma_i^{-1}\|$ (becoming more vulnerable to corrections from others). In a 1D space, this reduces to scalar variances σ_i and their reciprocal (i.e., the precision σ_i^{-1}). Authors publishing across diverse topics can capture many views (receive positive feedback) but, at the same time, they risk being corrected (receive negative feedback) when competing with other authors for viewership.

To consider player payoffs, start with an attack y at time t. Let y be a random independent sample on Y, and consider the problem of deciding whether the true distribution of Y is S_i or S_j. According to the Neyman-Pearson Lemma, the decision should be based on

$$\lambda^t(i, j) = \log\left\{\frac{S_i(y)}{S_j(y)}\right\}, \tag{6}$$

the likelihood ratio. Large values of $\lambda^t(i, j)$ favour the hypothesis associated with S_i and vice-versa. The ratio leads, in turn, to a family of optimal tests (O'Hagan and Forster 2004), each determined by a critical level b and the rule: decide in favour of i according to $\sum_{k=1}^{t} \lambda^k(i, j) > b$.

We take $\lambda^t(i, j)$ as the payoff for player i in a 2-player game at time t. Payoffs are thus derived from a relative measure of precision, and not the (incommensurate) absolute precision with which players detect the predator. This is in line with the notion that players are rewarded in proportion to how much they can correct (less accurate) others.

We next generalize this 2-player utility structure to a n-player game with a set $\mathcal{Y}^t = \{y_1, \dots y_t\}$ of previous attacks and $t \gg n$.

3.4 Equilibrium

The winner at time t is the player that spots the predator (at position y) without being corrected, as he can 'undercut' (correct the corrections of) all other players. So, we say that y is spotted by player $w_1(y)$,

$$w_1(y) = \underset{i}{\text{argmax}}\ \lambda^t(i, 0), \tag{7}$$

or, w_1 for short. And we say that the predator is spotted at that time with (absolute) precision $\lambda_0 = \lambda^t(w_1(y), 0)$.

Players gain from correcting others, but each player can only correct less precise others. The 'amount of correction' player $w_1(y)$ is guaranteed (above all other players) is then related to the precision of the second most precise player,

$$w_2(y) = \underset{j, j \neq w_1(y)}{\text{argmax}}\ \lambda^t(j, 0). \tag{8}$$

We say therefore that the winner $w_1(y)$ spotted the predator with (relative) precision $\lambda_1 = \lambda^t(w_1(y), w_2(y))$.

While $w_1(y)$ defines the winner, λ_0 and λ_1 define predator and player payoffs. Next, we imagine the predator gains by being spotted more erratically and a possible payoff for it is $-\lambda_0$.[4] On the other hand, a player $w_1(y)$ is guaranteed to be λ_1 more precise than all others. Thus, the game total payoff is the sum of predator and players' payoffs, for all spottings,

$$\Phi = -\log \prod_{y \in \mathcal{Y}^t} \frac{S_{w_1}(y)}{S_0(y)} + \log \prod_{y \in \mathcal{Y}^t} \frac{S_{w_1}(y)}{S_{w_2}(y)}, \tag{9a}$$

$$= \log \prod_{y \in \mathcal{Y}^t} \frac{S_0(y)}{S_{w_2}(y)}. \tag{9b}$$

Using the game as a prior, we will estimate players' distributions $S_i(y)$, $0 < i \leq n$, but not $S_0(y)$ directly. $S_0(y)$ is, however, analytically very significant. Equation 9 indicates that, collectively, players optimize a ratio between likelihood of attack, $S_0(y)$, and distance to others, $S_{w2}(y)$. We can get some insight about equilibria by reducing this game to a potential game with potential Φ.[5] In a potential game, the potential maxima are Nash Equilibria. And, in this case, it is the global maximum of Φ, indicating that precision increases incrementally with the addition of players.

3.5 A parametric PWM

As an analytic model, the PWM can be used in different ways, one possibility is introducing and estimating relevant parameters. With the assumption that players' areas are described by a Normal Distribution (with x_i as player i's position and Σ_i Covariance Matrix), each authorship or feedback event y carries information about the ratio

$$\frac{S_i(y)}{S_j(y)} = \frac{\mathcal{N}(y - x_i, \Sigma_i)}{\mathcal{N}(y - x_j, \Sigma_j)}. \tag{10}$$

where $w_1 = i$ and $w_2 = j$. We discuss alternatives for $\mathcal{D}_A^v$ - $x, y \in \mathcal{D}_A^v$ - across domains below. This type of representation and the consequent optimization problem, including Gaussian assumptions, appears sometimes across research fields (Moghaddam et al. 1998; Dong et al. 2018; Hillel and Weinshall 2007; Kostinger et al. 2012; Chen et al. 2019). Equation 10 is naturally associated with a distance function

$$d_i(y) = (y - x_i)^T \left[\Sigma_i^{-1} - \Sigma_j^{-1}\right](y - x_i). \tag{11}$$

[4] similar to Games Against Nature in Statistics (Blackwell and Girshick 1980) or Adversarial Training in Machine Learning (Goodfellow et al. 2018; Weinberger and Saul 2009), this stipulates that the best a player can do is to assume the other player will, also, do the best they can do. In the PWM, players organize to minimize risk of surprise (attacks) against a strategic opponent that, symmetrically, maximizes it. Such two-sided solutions often lead to robust strategies.

[5] The simple proof that Φ is a potential function is available in the Appendix, Sect. 1.

Learning this distance function (i.e., used by player i in this abstract game) corresponds to estimating the Covariance Matrices Σ_i and Σ_j.

We add a parameter β to this common representation which controls the mutual influence of content on players' decisions,

$$d_i(y) = (y - x_i)^T \Big[\Sigma_i^{-1} - \beta \Sigma_j^{-1}\Big](y - x_i), \tag{12}$$

when $\beta = 0$ players decisions are uncorrelated and when $\beta > 0$ content from other players have a deterring influence on players' topic decisions. The metric is associated, in turn, with the revised ratio

$$\log \frac{S_{w_1}(y)}{S_{w_2}(y)^{\beta}}. \tag{13}$$

In the PWM, the parameter is interpreted as the ratio, and relative significance, of predator attacks (content views) and other players (content corrections). After discussing document representation and metric learning in further detail, we introduce a Bayesian framework that can estimate β and Σ_i simultaneously. Estimating a Maximum a posteriori (MAP) likelihood β and Covariance Matrices Σ_i across players leads to a shared MAP metric. Finally, we will use the resulting individual player distributions to predict user behavior and β as a parameter to describe topic dispersion or contagion in social media.

4 Game and document representation

To simplify and scale the approach to large corpora, we assume a simple document representation and topic-word distribution, $P(w|y_k)$. Table 1 summarizes how documents are represented in each of the studied domains, namely, as bags-of-verbs or bags-of-ingredients.

Consider, for example, a set A of known verbs of size v. A document k can be represented by a position y_k in the simplex of actions $\mathcal{D}_A^v$, having coordinates

$$y_k(a) = \frac{c_k(a)}{\sum_{a \in A} c_k(a)}, \tag{14}$$

with c_k a v-sized vector where coordinate $a \in A$ has the number of times k mentions a. For Wikipedia, these coordinates then represent a page describing a real-world entity and its possible 'actions' (taken altogether, the entity behavior), with $P(a|y_k) = y_k(a)$. Since actions are simplex corners, this also puts actions and players in a common space. This is a simple but common representation in game theory.

We can now consider user i's set of authored documents, $U_i \subset \mathcal{D}_A^v$. We have defined a user's position as the mean position over these documents,

$$x_i(a) = \frac{1}{|U_i|} \sum_{y_k \in U_i} y_k(a). \tag{15}$$

We thus define the Mahalanobis distance between a user i and document k as

$$d_i(x_i, y_k) = (x_i - y_k)^T \Sigma_i (x_i - y_k), \tag{16}$$

where Σ_i are $v \times v$ Covariance Matrices to be learned. We denoted $\mathcal{D}^v_A$, in particular, the metric space with dimensions from A. This definition makes the optimizing metrics space, $\mathcal{D}^v_A$, the cone of symmetric positive definite $v \times v$ real-valued matrices (Bar-Hillel et al. 2005). The use of Covariance Matrices to define and learn distance metrics is commonplace in Metric Learning research (Bellet et al. 2013).

With the previous elements, we can formulate the final stochastic model employed in this article as

$$\begin{aligned}
&P_i(a) = P(i) \int\limits_{y_k \in \mathcal{D}^v_A} P(a|y_k)P(y_k|i),\\
&P(y_k|i) \sim \mathcal{N}(x_i, \Sigma_i), \quad (user - document)\\
&P(a|y_k), \quad y_k \in \mathcal{D}^v_A, \quad (document - word)
\end{aligned} \tag{17}$$

where document-word distributions, $P(a|y_k)$, are defined over actions in A and user distributions, $P(y_k|i)$, over documents.

5 Metric learning

Metric Learning is formulated in this article largely as the problem of estimating β. The parameter was formulated in the PWM from a relationship among players' Covariance Matrices. In the Mahalanobis Metric Learning literature (Bishop 2006; Goldberger et al. 2004; Kostinger et al. 2012), learning is often unsupervised. It proceeds by assembling a Covariance Matrix from pairwise differences over data points, which is then inverted to provide the metric. Here, the metric does not follow deterministically from a single player Covariance Matrix, as it depends on the parameter β that relates all players' Covariance Matrices. The parameter β will carry, in fact, most of the predictive power in the trained models. It is seen as a property of the medium and serves to both predict and describe behavior in the studied medium from its players' Covariance Matrices. The overall training procedure will consist therefore of assembling player Covariance Matrices, and then, finding a Maximum A Posteriori (MAP) estimate for β that is predictive of players' behavior, collectively.

5.1 Player metrics

A Mahalanobis metric associates a high-dimensional central position with a Gaussian mean and distances from that position with a Covariance Matrix. Since the position x_i for a fixed player i is noisily observed (from her past documents), the main goal of inference is her distance to all other players, x_j, and all content, y. Fixing the player leads to a player-specific metric (all player metrics related by β). We take x_i and β as given in this section and consider only two players. We return to the estimation of x_i and β, and all players, in the next section. Because, in each player-specific metric, x_i is fixed, it becomes convenient to express Eq. 13 as a ratio from the mean. This is a common strategy in Mahalanobis Metric Learning (Hillel and Weinshall 2007; Kostinger et al. 2012; Chen et al. 2019) - as it allows the metric to be formulated exclusively with operations over Covariance Matrices. In this case, each player distribution is centered at the mean, $\mathcal{N}(0, M_i)$, where M_i has been biased to reflect the fixed mean.

Consider then the displacement vector between player i's position, x_i, and her documents' positions $y_k \in U_i$, $y_k - x_i$. Represent this vector set with a zero-mean Gaussian $\mathcal{N}(0, M_i)$ where the Covariance Matrix is

$$M_i = \frac{1}{|U_i|} \sum_{y_k \in U_i} (y_k - x_i)\,(y_k - x_i)^T\,. \tag{18}$$

The matrix is defined similarly for other players.

Assume that $w_1(y_k) = i$ and $w_2(y_k) = j$ are the two nearest users to y_k. Given the PWM, the first is known since player i authored y_k. The player j is only known a priori in a two-player game. According to the model, Eq. 13, player i's strategy in this case is to maximize the ratio

$$\log \left\{ \frac{\mathcal{N}(0, M_i)}{\mathcal{N}(0, M_j)^{\beta}} \right\}. \tag{19}$$

For multiple players, the problem becomes that of changing the metric such that the ratio is increasingly larger when w_2 is also unknown. The accumulated metric change from millions of players is achieved here by random sampling distinct players and adjusting the metric with each pair. This step is spelled out in the next section. Before that, we consider the pairwise (i.e., two-player) case, whose logic follows closely recent work in large-scale Mahalanobis Metric Learning (Hillel and Weinshall 2007; Kostinger et al. 2012; Chen et al. 2019).

For player i and j, Eq. 19 is

$$\begin{aligned}
&\log \left\{ \frac{\mathcal{N}(0, M_i)}{\mathcal{N}(0, M_j)^{\beta}} \right\}, \\
&= \log \mathcal{N}(0, M_i) - \beta \log \mathcal{N}(0, M_j), \\
&= \log \frac{\exp\Big(-\frac{1}{2}(y_k - x_i)^T M_i^{-1}(y_k - x_i)\Big)}{\sqrt{2\pi}|M_i|} \\
&\quad -\beta \log \frac{\exp\Big(-\frac{1}{2}(y_k - x_i)^T M_j^{-1}(y_k - x_i)\Big)}{\sqrt{2\pi}|M_j|},
\end{aligned} \tag{20a}$$

$$\begin{aligned}
&= (y_k - x_i)^T M_i^{-1}(y_k - x_i) + \log |M_i| \\
&\quad -(y_k - x_i)^T \beta M_j^{-1}(y_k - x_i) + \beta \log |M_j| + Const.,
\end{aligned} \tag{20b}$$

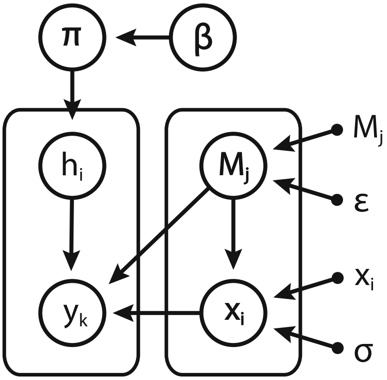


Fig. 2 Graphical Model for the estimation of PWM parameters as a Hierarchical Gaussian Mixture, hierarchical models consider possible errors in the measuring of model variables, posterior variables are bolded, for player i, $\mathbf{x_i}$ and $\mathbf{y_k}$ are posterior player and document positions, $\mathbf{M_j}$ are player Covariance Matrices, $\boldsymbol{\beta}$ is a medium parameter, $\boldsymbol{\pi}$ is a mixture weight, h_i is an author-document indicator and M_j, x_i, ξ, σ (non-bold) are priors for the previous variables, assuming they follow Wishart and Normal distributions

since the log and constant terms in Eq. 20b are offsets (Hillel and Weinshall 2007; Kostinger et al. 2012; Chen et al. 2019) for a given player, they are ignored. The Mahalanobis metric that maximizes the previous log-ratio test (and also corresponds to the least-squares Gaussian Maximum Likelihood estimate from the mean, given β) is then

$$d_i^{\beta}(x_i, y_k) = (y_k - x_i)^T (M_i^{-1} - \beta M_j^{-1})(y_k - x_i), \tag{21}$$

and, thus, we make $\Sigma_i = (M_i^{-1} - \beta M_j^{-1})$. We can finally formulate player i's distribution (considering the influence of other players) as

$$S_i(y; \beta) = \mathcal{N}(x_i, \Sigma_i). \tag{22}$$

5.2 Parameter estimation

We now describe a Bayesian Maximum a Posteriori (MAP) estimation procedure for β and Covariance Matrices Σ_i. Beyond generalizing the previous two-player case, this has the extended benefit of modeling measurement errors in the estimation of player positions x_i, matrices Σ_i and parameter β (Steinberg et al. 2015). We assign them Gaussian, Wishart and Uniform priors respectively,

$$\begin{aligned} \mathbf{x_i} &\sim \mathcal{N}(x_i, \sigma(M_i{}^{-1} - \boldsymbol{\beta}\mathbf{M_j}{}^{-1})), \\ \boldsymbol{\beta} &\sim U([-1, +1]), \\ \mathbf{M_j} &\sim \mathcal{W}(M_j, \xi), \\ \boldsymbol{\pi} &\sim Dir(\beta), \end{aligned} \tag{23}$$

where σ and ξ are error parameters (posterior variables bolded) and $U([-1, +1])$ is the unidimensional uniform distribution in the interval $[-1, +1]$.

We considered that all authored documents $y_k \in \mathcal{D}_A^v$ are drawn from a mixture of n author distributions,

$$P(y_k) \sim \sum_{i=0}^{n} \pi_i \mathcal{N}(y|\mathbf{x_i}, \Sigma_i), \tag{24}$$

where $\boldsymbol{\pi} = \pi_{i=1}^n$ are mixture weights, $\pi_i \in [0, 1]$, and $\sum_{i=1}^n \pi_i = 1$. Correspondingly, $\mathbf{x_i}$ and ˝$_{\mathbf{i}}$ are the posterior means and inverse Covariance Matrices, ˝$_{\mathbf{i}} = \sigma({M_i}^{-1} - \boldsymbol{\beta}{\mathbf{M_j}}^{-1})$, for each author.

As before, we call i a document's author and j opponents (other players). While player $w_1 = i$ is known a priori (who authored or fedback a document), we make the assignment of player w_2 random, with the probability that a player j is $w_2 = j$ following the estimated mixture weights $\boldsymbol{\pi}$ and a categorical distribution.

The detailed implementation for this estimation procedure is based on a common Hierarchical Bayesian Gaussian Mixture Model. It is the simplest such hierarchical model in the range currently in use for image or cluster models (Steinberg et al. 2015). Steinberg (Steinberg et al. 2015) provides a good summary for these models and techniques. Images are in the present case replaced by the document representation defined by Eq. 18. The resulting procedure continuously draw opponents $w_2 = j$ that can serve as evidence to adjust parameters in a Bayesian fashion. Accordingly, the matrix M_i is fixed for each author, Eq. 18, while M_j is estimated iteratively. For large-scale media with thousands or millions of authors is reasonable to assume that M_j is the same for all authors. M_i expresses authors' specialization or deviation from M_j, which, in turn, expresses common or shared knowledge.

The graphical model of the process is shown in Fig. 2. We denoted f_k the feedback received by document k, which is an integral count. The counts used in experiments are summarized in Table 1 (counts of views, edit reversals, revenue dollars and yums). It is common (Steinberg et al. 2015) to also introduce an auxiliary indicator variable, $H = h_{k=1}^K$, where $h_k \in \{1, ..., n\}$ and K is the number of documents used for training. The variable h_k randomly assigns an opponent w_2 to each training observation (i.e., an alternative author j), and thus Gaussian component, leading to the conditional relationship

$$y_k|h_k \sim \prod_{j=0}^{n} \mathcal{N}(y_k|\mathbf{x_i}, \Sigma_i)^{\mathbf{1}[h_k=j]}, \tag{25}$$

where $\mathbf{1}[.]$ takes value 1 when the bracketed expression is true, and 0 otherwise. Both sampled documents k and opponents h_k are distributed according to Categorical distributions,

$$\begin{aligned} k &\sim Categ\left(f_k / \sum_{v=0}^{K} f_v\right), \\ h_k &\sim Categ(\boldsymbol{\pi}) = \prod_{j=0}^{n} \pi_j^{\mathbf{1}[h_k=j]}. \end{aligned} \tag{26}$$

This is justified by heterogeneity in players' influence across social media, which is the central concept in many social media models (Goyal et al. 2010; Cosley et al. 2010; Demarzo et al. 2003).

A useful way of thinking (Bishop 2006; Steinberg et al. 2015) about Bayesian Mixture Models is to imagine that each data point y is associated with a latent indicator variable $h_k \in \{1, ..., n\}$ specifying which mixture component generated that data point. These assignments are analogous to class labels in a Bayesian classifier, except that they are now stochastic. In the present case, labels correspond to opponents. The pairwise model in Eq. 21 is then used to establish a generative relationship between the author, selected opponent and observation y_k. We write, as a result, the final published content log-likelihood as the product of the indicator distribution, $Categ(\boldsymbol{\pi})$, and the authorship distribution, $\mathcal{N}(y_k | \mathbf{x_i}, \Sigma_i)$, which is what Eq. 25 expresses. The central outcome is a MAP estimate for β and Covariance Matrix Σ_i which are used for authorship prediction.

The resulting sampling procedure is implemented with variational Bayes (Steinberg et al. 2015; Attias 1999) and is summarized as follows:

1. Draw K documents $k \sim Categ(f_k / \sum_{v=0}^{K} f_v)$.
2. Draw β and mixture weights $\boldsymbol{\pi} \sim Dir(\beta)$.
3. For each training document, $k \in 1, ..., K$,
   (a) Draw an alternative author $h_k \sim Categ(\boldsymbol{\pi})$.
   (b) Draw author parameters $\mathbf{x_i}$, Σ_i, with $y_k | (h_k = j) \sim \mathcal{N}(\mathbf{x_i}, \Sigma_i)$.

6 Complexity and implementation

The previous particular solution and representation favors scalability. First, the set of verbs is typically limited (and has simpler and more 'stable' meanings than those in names and entities collections). The Covariance Matrices are all square v, not n (e.g., v is 2 orders of magnitude smaller in the studied domains). Representing players as sparse vectors over this limited lexicon curtails memory requirements with content growth (i.e., more documents). Second, metric learning (Eq. 21), and the matrix inversions it requires, can be implemented with a Cholesky-Decomposition (Bar-Hillel et al. 2005; Kostinger et al. 2012), avoiding the complexity of a full-SVD or other more complex computations typical in Metric Learning (Weinberger and Saul 2009).

We can now consider how to calculate the vectorial representation for each document. This amounts to simple verb or ingredient frequency counts, Eq. 14, for each document in a medium. Each user is then associated with a document and vector set U_i.

To parse verbs, we used (sen 2014) for sentence boundary and (Collobert et al. 2011) for semantic role labeling of text fragments, both containing good quality/complexity trade-offs. An inverted document index from the set of all strings with semantic roles of action to documents is built from the labeler output. The resulting verb lexicon A is also transformed to root form with a morphological dictionary (xta 2014). The index allows the efficient calculation of document-verb counts and user-document indicators. With these, we can then calculate Eq. 14 for each document and Eq. 15 for each user.

All documents are processed chronologically. For Wikipedia, pre-processing involves a further Wiki-text parser and the inverted index used is a pre-trained link-resolver dataset (Singh et al. 2012) that outputs Wikipedia page titles. For movie scripts, dialogue text is ignored. For recipes, ingredients are listed in separate, and easily parsed, document sections and no lexical labeling is necessary. We discuss further problem-specific details in the experimental section.

With the document representation, U_i, we can calculate user positions x_i (Eq. 14, 15), followed by matrices M_i for each player (Eq. 20b). The final player distributions (Eq. 22) can be easily obtained from these matrices after parameter estimation. The training procedure is therefore summarized by the following steps

Input: Set of k Authored or Shared Documents from each of n users, $0 < i \leq n$, and their feedback counts f_k.
Output: Set of Distributions for each user, $S_i = \mathcal{N}(x_i, \Sigma_i)$.

1. Transform user documents to vectorial form, U_i, using Semantic Role Extraction and word frequency counts, $0 < i \leq n$,
2. Calculate the medium's β and matrices Σ_i using the Hierarchical Bayesian Model,
3. Calculate individual user distributions, $S_i = \mathcal{N}(x_i, \Sigma_i)$, consisting of user positions, x_i, and Covariance matrices, Σ_i.

7 Experiments

All datasets consist of documents labeled with author ids, feedback counts and time stamps. They exemplify different uses of media. In all cases, feedback variables f_k are counts (feedback aggregates), which are commonly found across social and general media datasets. We start with authorship prediction, then turn to feedback prediction.

We compare performance of the model **by reproducing multiple previous studies and extending them with the proposed method and state-of-the-art topic model methods** such as the Dirichlet-Multinomial Regression (DMR) (Mimno and McCallum 2008). In particular, we consider accuracy gains brought by the devised metric, when we add it (or fully replace) the feature sets of different studies. We will review the original experimental protocols but more details are available in the original publications in each case.

7.1 Wikipedia

We start with authorship in Wikipedia from its inception to current days. We use the PWM to predict user authoring behavior - author given documents or editing behavior given author. Wikipedia is a public corpus and has been used in previous studies (see below). We pay special attention to *author* behavior prediction in this case (which topics and how much users will publish over in the future). We take documents to be Wikipedia pages (a person, country, etc.) and authors to be page editors. All results use the May 1st 2015 Wikipedia dump. This involves approximately 2.9M pages, 11M (registered) users, 200k verbs and 700M edit reversals. Edit reversals are Wikipedia edits that are not accepted. We then calculate player distributions (Eq. 22) for every user.

To calculate these distributions, we first carry out a MAP estimation of β and M_j (Sect. 5.2). For convenience, we consider a Wikipedia parameter training subsample, consisting of the 300k most viewed pages (Wikimedia 2012) and all their editors (including authors) until 01/01/2010.

Before evaluating prediction performance, let's consider the resultant distance metric. Consider a given document y and a document author i, define **author-to-doc** and **other-to-doc** distances to be $d_i(x_i, y)$ and $d_i(x_{\hat{w}_2}, y)$, where w_2 is the closest user to y among all other users,

$$\hat{w}_2 = \underset{j, j \neq i}{\operatorname{argmin}}\, d_j^{\beta}(x_j, y), \tag{27}$$

according to the trained model. The PWM suggests that users will edit documents both close to them and distant to others. Informally, these correspond to documents that are, simultaneously, of interest to the user and 'unattended'. The ratio between these two distances correspond to the ratio in Eq. 9a when $w_1 = i$ and $w_2 = j$. Figure 3a shows a histogram with counts of author-to-doc and other-to-doc combinations for all Wikipedia page edits. Each cell counts edits with a given combination (discretized in 50-by-50 subranges). It illustrates Wikipedia users' propensity to edit documents that are close to their positions (his or hers past documents) while simultaneously distant to other authors. Notice that Wikipedia pages have multiple authors and this is a stochastic relation.

Wikipedia dumps (and Wikipedia itself through its user interface) provide authorship information. Audience behavior is, however, also available from a Wikimedia site, as a dataset of Wikipedia page views (Wikimedia 2012). We can repeat the previous procedure using, instead, audience data and positive feedback (page views). The

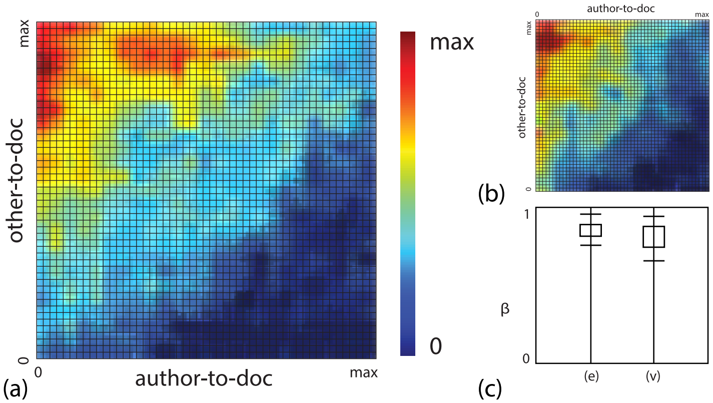


Fig. 3 **a** User to document distances for every Wikipedia edit, each cell in the 50 × 50 histogram counts the number of edits with a given combination of author-to-doc and other-to-doc distances (see text for definitions), all edits observed in Wikipedia until 01/01/2016, users edit pages that are both close to their previous edits and distant from others' edits, **b** repeated results using page view data (instead of edit reversals), **c** MAP β estimates for Wikipedia with (e) edit reversals and (v) page view data

existence of an equilibrium between authors and audience imply parameters and distances should be the same when using feedback from authors or audience.[6] Figure 3b reproduces the previous histogram in this case and Fig. 3c shows β estimates in the two cases.

Although we focus on behavior prediction in this article, parameters can also be used to help describe and understand behavior in the medium through time. Figure 4a shows the increase in estimated β with Wikipedia's increasing author base. This suggests how the increase in author numbers leads to increased incentives for 'specialization' (authors' tendency to publish over increasingly similar documents).

7.1.1 Wikipedia challenge

We are aware of four results that outperform pure contagion models (Cosley et al. 2010) when predicting Wikipedia user behavior: two Wikipedia Participation Challenge entries, recent topic models (Cha and Cho 2012; Rosen-Zvi et al. 2004) and a recent Graphical Model (Zhang et al. 2014).

We first consider the 'Wikipedia Participation Challenge' sponsored by Wikipedia itself (the Wikimedia foundation), it 'challenge[d] participants... to predict the number of edits an editor will make five months from the end date of the training dataset' (wik 2012). The challenge grew out of a practical social media problem, that 'between 2005 and 2007, newbies started having real trouble successfully joining the Wikimedia community... the community ha[s] become too hard to penetrate...' The challenge made available a >3GB dataset of randomly sampled pages and edits from the English Wikipedia from January 2001 to August 2010. But participants were also allowed to use any pre-September 2010 data from Wikipedia dumps. Accuracy was measured

[6] In Eq. 9, distributions S_0 (audience views) and S_{w2} (author corrections) are symmetric except for the sign, leading to similar Eq. 19.

Table 3 Wikipedia Participation Challenge: Will a given user edit any pages?

	Prediction Error (RMSLE)							
Classifier/Feature-set	Top-1	Top-2	PWM	DMR	Top1 +PWM	Top2 +PWM	Top1+DMR	Top2+DMR
Logistic (wik 2012)	0.79	0.85	0.4003	0.97	0.3145	0.3099	0.4179	0.3991
Random Forest (wik 2012)	0.86	0.84	0.3647	0.9576	0.29184	**0.27786**	0.3923	0.3719
(perplexity score)	–	–	0.4098	1.1281	–	–	–	–

in Root Mean Squared Logarithmic Error ('RMSLE'), with the Wikimedia in-house solution starting at 1.47708, and the final winner reaching 0.791274 (Table 3).

For any user, the first task asks whether users will remain active Wikipedia contributors (edit any pages 5 months in the future). No submitted solution attempted to use (natural language) edit/page content, with the top two training (resp., linear and random-forest) binary regressors for users (with outputs: will-edit and not). The first used 13 features (2 based on reverts and 11 on past editing behavior) and the second 206 (with edit timing and editing volume deemed most informative). Description and code are open-source (wik 2012).

Since the challenge involves per-user binary prediction, we trained the same regressors as the top-two entries but used Eq. 22 as single feature (PWM) or as an additional feature (Top1+PWM and Top2+PWM) to the previous best-performing solutions. Top1 is the set of features used by the challenge's winner and Top2 by the runner-up. We also repeated this procedure using the probability of an unseen document (page) given a user as given by a Dirichlet-Multinomial Regression (DMR) topic model (for details, see the next subsection). The PWM and DMR features suggest the relative predictive power of content (text). A Random Forest classifier ($n = 200$) with added PWM features generates the lowest RMSLE, 0.27786 (Table 3). The RMSLE in this case is

$$\epsilon = \sqrt{\frac{1}{\hat{n}}\sum_{i=1}^{\hat{n}}(\log(p_i + 1) - \log(a_i + 1))^2}, \tag{28}$$

where $\hat{n}$ is the total number of users in the test data set, p_i is the regressor output, and a_i is the actual response for user i.

In addition to identifying the best performing algorithm, this also indicates how the different features contribute to prediction. Performance deteriorates when any of the text-based features are removed; indicating that content can contribute to user behavior prediction. Measured by RMSLE, PWM features are the biggest single contributor as the regressors without them see the greatest deterioration, followed by DMR. The single-feature PWM reaches RMSLE close to the best (Top2+PWM) and gains less than the DMR from other features. This demonstrates that the PWM incorporates global social media information succinctly into its authorship predictions.

7.1.2 User graphical and topic models

The other behavior prediction results highlight another side of this problem, predicting user interests (i.e., which *individual* pages a user is likely to edit). Since we took Wikipedia pages as possible documents, Eq. 22 predicts interest directly (no regression needed). We start with the experiment and solution proposed by Zhang et al. (2014). They define a bipartite user-page graph and use the sum-product algorithm. They studied different graphs. Graph edges were derived from Wikipedia users' social networks and pages' subject categories (thus making use of the Encyclopedia-like ontological relationships curated in Wikipedia). Authors report F1-Accuracy of their solution, 0.87, and the standard content-based recommendation system of Segaran (2007), 0.37. They use a subset of 700 editors in the Wikipedia Category C1 = 'Wikipedians interested in art'. PWM scores a F1-Accuracy of 0.912 on (the same) 3-months interval (Table 4). We disregard Wikipedia 'social' or category data but use page text.

We also performed the task with subsequent semesters of Wikipedia data, Fig. 4b. Scores suggest how β estimates converge with sequential edits. Figure 4c shows performance with a percentage of largest-area only users (i.e., with area-ordered and increasing n in Eq. 20b). It suggests that most performance can be ascribed to the 10th percentile area users.

The Dirichlet-Multinomial Regression (Mimno and McCallum 2008) extends LDA to take meta-information (e.g., document features). Author information is typically introduced with binary author indicator features (Mimno and McCallum 2008). The model's ability to predict authors of a held-out document conditioned on the words ('author prediction') have shown to outperform significantly the Author-Topic Model (ATM) (Mimno and McCallum 2008; Rosen-Zvi et al. 2004). This is done by defining a non-author-specific Dirichlet prior on topics (i.e., the prior for a document with no observed features).

We use DMR on the experiment of (Zhang et al. 2014) and extend it with other four Wikipedia categories (using the same protocol and train/test timeline). The categories are C2 = 'Wikipedians with PhD degrees' ($n = 1,073$), C3 = 'American Wikipedians' ($n = 4,396$), C4 = 'Wikipedians in England' ($n = 1,568$), C5 = 'Wikipedians in Canada' ($n = 1,409$). We report DMR performance with author indicators (Mimno and McCallum 2008) or distance-to-doc (as given by the PWM). In the first (DMR), feature vectors (of size n) are binary and indicate whether the author edited the document/page. In the second (DMR+PWM), feature vectors contain author-document distances calculated with Eq. 22. We use 100 topics and optimizations as in Mimno and McCallum (2008). The results, Table 4, are reminiscent of the previous edit-or-not study, indicating that the PWM prior 'social' model carries predictive power. Table 4 (bottom row) additionally shows perplexity scores often reported for LDA solutions, which suggests that the model can also help ameliorate the data requirement of generative models (an often-noted downside). The PWM without any generative component, on the other hand, gives a simple and scalable alternative to LDA in this dataset, requiring only the estimation of a medium wide parameter β that characterizes user behavior in the medium. Knowing β for Wikipedia, in particular, allows author behavior prediction with simple matrix operations.

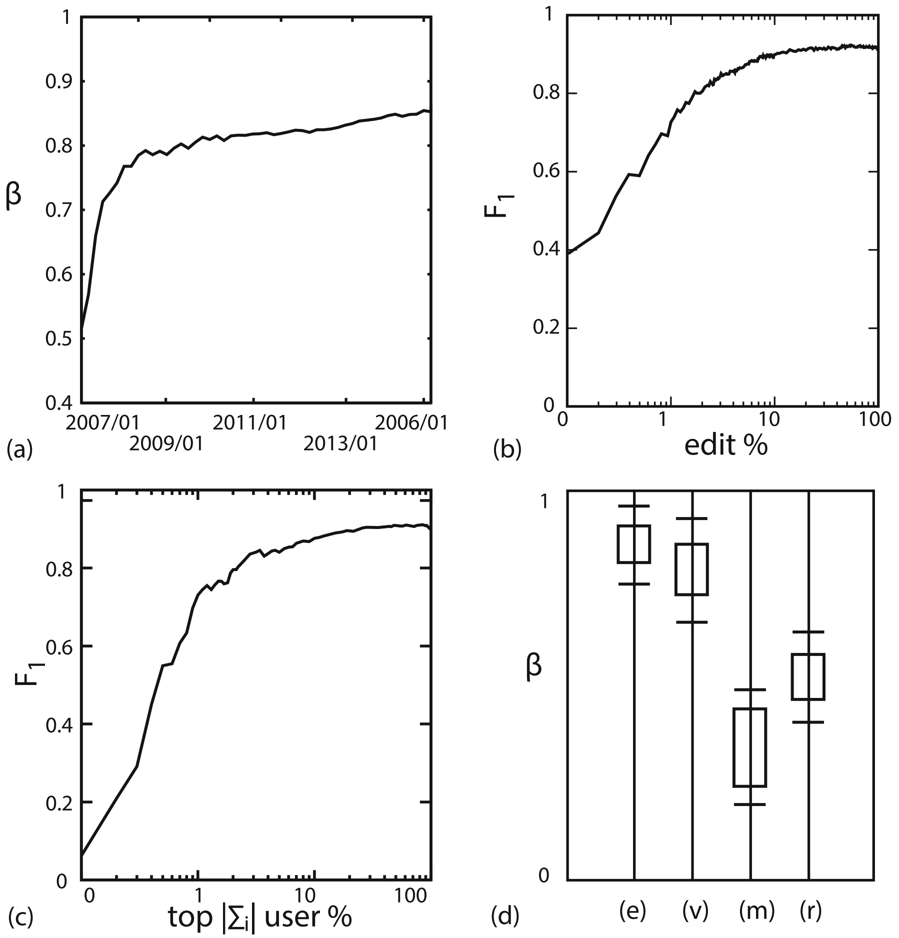


Fig. 4 **a** Parameter β estimated across Wikipedia history (2007-16), dispersion grew until 2008 when it reached a tableaux with slower growth, **b** PWM accuracy versus % edits taken chronologically (abscissa in log-scale), (**c**) PWM accuracy versus percentage of documents ordered by $|\Sigma_k|$ (abscissae in log-scale), (**d**) MAP β estimates for (e) Wikipedia using edit reversal data, (v) Wikipedia using page view data, (m) movie scripts using box-office revenue, (r) food and drink recipes using yum counts; while all media are dispersive, Wikipedia is the most topically dispersive medium

7.2 Movie scripts and recipes

We considered Wikipedia user behavior (authorship) prediction in the previous section. We now consider a second large-scale authorship scenario where authors are simultaneously aware of other authors and their audience. The problem of predicting the success of movies has been studied extensively (Lash and Zhao 2016; Ghiassi et al. 2015; Eliashberg et al. 2014). The most typical features used have been pre-release movie features such as movie genre, actors, audience demographics and production values. More recently, attention has turned to textual data such as plot summaries, reviews and scripts.

Table 4 Wikipedia User Behavior Prediction: Which pages a given user will edit?

	Accuracy (F1)				
Algorithm/Experiment	C1	C2	C3	C4	C5
CF (Segaran 2007)	0.37	–	–	–	–
Sum-Product (Zhang et al. 2014)	0.87	–	–	–	–
PWM	0.944	0.8789	0.9389	**0.9211**	**0.9632**
DMR	0.761	0.5899	0.8010	0.7599	0.8118
DMR+PWM	**0.942**	**0.8989**	**0.9579**	0.8900	0.9576

Table 5 Movie-Scripts Box-office prediction: Will audiences pay to watch a movie?

Feature-set/Regressor	Mean-Square Error (MSE)	
	Bayesian Additive Regression Tree (BART)	Kernel-II
Budget+Structure+LSA (Eliashberg et al. 2014)	0.5342	0.4219
Budget+Structure+DMLR	0.4012	0.3712
Budget+Structure+DMLRA	0.3989	0.3638
Budget+Structure+PWM	**0.3104**	**0.2791**

We reproduce the study of Eliashberg et al. (2014), who are the first to use the full text of scripts. We take Authors to be directors in this case. The study uses four classes of features to predict box-office performance, which the authors call: genre and 'content' variables, structure variables, LSA variables and the production budget (in dollars). The first set of variables is obtained from questionnaires administered to independent readers. Responders were asked to identify important aspects of scripts often used by screen writing experts (e.g., movie genre, clear premise, believable ending). Structure variables describe the script structure[7]: total number of scenes, percentage of interior scenes, total number of dialogues, average number of dialogues and a concentration index of dialogues. LSA variables are Latent Semantic Analysis (LSA) positions, calculated after stemming and removing low TD-IDF words and stopwords from scripts. The study uses log-box office revenue (in $ Millions) to measure movie performance.

We ignore the first subjective class of variables and extend the full set of scripts to a total of 4085 from the 300 used in the original study. Figure 5a shows distances combinations for all documents with a 13 × 13 histogram. It repeats the pattern seen in Wikipedia and formulated by the PWM. We also add release date and authorship features, which we will take to be the movie's director. We consider the first because the increase in sample size expands the date range from 3 years to over 40 years. We consider the second to capture information about directorial specialization and the competitive nature of movie making. This simply consists of a unique identifier for

[7] The authors originally called these 'semantic' variables.

Table 6 Movie-Scripts Box-office prediction: Feature Imputation

	Mean-Square Error (MSE)	
Feature-set/Regressor	Bayesian Additive Regression Tree (BART)	Kernel-II
Budget+PWM	0.3501	0.3121
Structure+PWM	0.3021	0.3005
PWM	**0.3191**	**0.2925**
DRMLA	0.4172	0.3977

the director. We then repeat the study's hold-out procedure but use 360 test movies released after 2009, instead of only 30.

The original article reports box-office revenue forecasts with Mean Squared Error (MSE) of approximately 40%. Authors conclude that such results hold great promise because they rely solely upon data available in pre-production (the scripts themselves). They consider multiple regression-based solutions and a proposed kernel-based solution. We consider their two methods with best performance: a Bayesian additive Regression tree (BART) and their Kernel-II ('optimized' feature weights). We consider DMR and PWM features as alternative to the LSA-based features in the original study. Notice that by adding PWM features for the director, we are making predictions for the success of a director on a set of produced movies. LSA assumes a latent Euclidean semantic space from word frequencies, and take documents to be positions in the space. Due to this, LSA has the attractive feature that document positions and their distances are generally interpretable. LSA methods are however often less competitive compared to LDA-based methods, such as the DMR. This is confirmed in Table 5. DMR denotes a standard DMR model and DMR+A a DMR model with author indicator features (similar to those described in the Wikipedia task). Regressors with DMR features outperforms the original LSA representation, but not the PWM which outperform both, bringing the MSE down to 0.279.

Maybe surprising, this brings revenue prediction below the error from approaches using the highly curated and subjective 25 'content' features. We next consider to what extent this increase in improvement is due to the two added features (release date and director identifier) or the sample size increase. Results are in Table 6. Notice that the addition of authorship and time features don't lead, by themselves, to significant increases in accuracy. These variables, however, do lead to increased performance when combined with prior statistical models, either purely semantic (such as the DMR) or strategic (such as the PWM). The PWM lead to even increased accuracy gains in the case of movie scripts.

We consider one last scenario, which hasn't been considered previously in the literature. It demonstrates performance in an example where the semantic and lexical challenges are simpler. Wikipedia pages and movie scripts are typically medium to large length documents and consist mostly of free-form texts. Recipes are relatively small, with more structured content. Instead of the worldly characters in Wikipedia or the idealized characters in movie scripts, cooking recipes are mostly about ingredient combinations. We thus take documents to be combinations of ingredients in this

Table 7 Food and Drink Recipes feedback prediction: Will audiences yum?

	Mean-Square Error (MSE)	
Feature-set/Regressor	Bayesian Additive Regression Tree (BART)	Kernel-II
LSA	0.7321	0.7176
DMRA	0.5611	0.5872
PWM	**0.1644**	**0.1578**

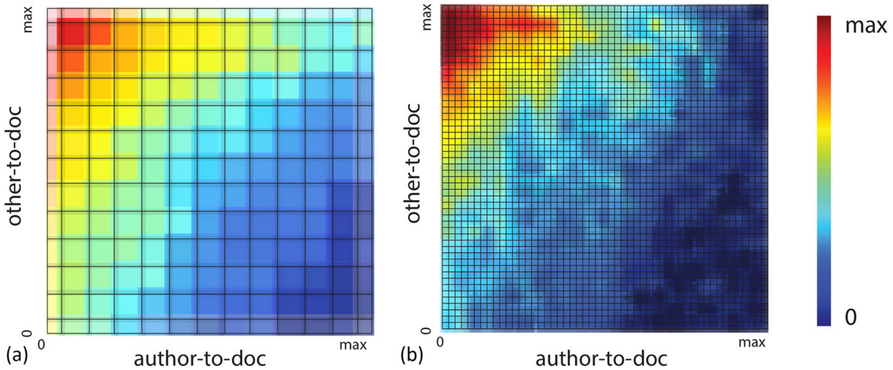


Fig. 5 User to document distances for every (**a**) movie script (13 × 13 histogram) and (**b**) food and drink recipe (50 × 50), each cell counts the number of documents with a combination of author-to-doc and other-to-doc distances for all author and document pairs; this illustrates a common authorship pattern in the two media, also observed for Wikipedia

example. Like the set of verbs considered in Wikipedia and movie scripts, the set of all ingredients are constrained and known.

Similar to movie revenue prediction, we predict in this case the number of yums (positive feedback) that a new recipe gets in the popular recipe-sharing site yummly.com. The dataset consists of 369 thousand drink and food recipes in yummly from its creation in 2010 to 2013. Figure 5b shows distance combinations in this dataset. We take recipes in the last 6 months as the testing sample. As before, each training point consisted of author, content (ingredients) and time stamps. Results are in Table 7. In this case, purely semantic (User-Content) models perform much worse than the PWM. Results in this simpler dataset indicate that strategic authorship and content models can be combined and are complementary. Their combination could lead to solutions that are accurate across wider ranges of content types.

7.3 Parameters and future work

Figure 4d shows the MAP estimate for β for the studied domains, where the null-hypothesis $\beta = 0$ would indicate no mutual adaptation among users (making Σ_i in Eq. 22 the sample covariance over the documents they author). This indicates that past individual data alone, without a prior social behavior model, might be insufficient to predict authors' behavior. Instead of looking at the problem locally and asking whether a user will share a given post, framing authorship prediction collectively, for all users, can reveal mutual adaptations and regularities.

In this article we presented a model for online authorship and resulting techniques. These techniques can be employed in the study and prediction of authorship within and across media. This is because the stipulated model elements appear across many media (in particular, the reciprocal exchange of text content and feedback). It thus has direct applications in problems that require user and consumer predictions (such as recommending systems, business support, social media design, etc.) The error estimates in Fig. 4d suggest that parameters such as β could characterize media and the authoring behavior observed in them, where the behavior observed in Wikipedia (topic 'dispersion'), or possibly the opposite (e.g., $\beta = -1$ and 'topic swarming'), could be observed in other popular media like Facebook or Tumblr. Understanding what affects and changes authorship patterns across media (and its relationship to designed feedback mechanisms) is a largely unexplored area for research. The framework developed here suggests one way to study the effect of such design issues on content authorship and sharing.

8 Conclusion

We proposed techniques to study and predict how people adapt their authoring behavior in view of shared content and each other. We assumed that, whenever networked together, people develop mutually recognized roles and expertises that shape their behavior. Beyond techniques, we introduced a game describing the strategic problem and content-based behavior prediction results when content exchanged is encyclopedic knowledge, movies or food recipes.

Funding This work was supported by a Cambridge Overseas Trust Fellowship (University of Cambridge).

Availability of data and material (data transparency) Not applicable.

Declarations

Conflict of interest The author declares no conflict of interest.

Code availability (software application or custom code) Not applicable.

Appendix

The PWM is a potential game

We defined the total payoff in the game as the sum of predator and players' payoffs, Eq. 9. At most one player is getting any payoff from a given spotting; this suggests the possibility of a potential in the game. To show the game has a potential, we need to show that Φ reflects additively a player's payoff change when he leaves or enters the game (Milchtaich 1996). Consider then what happens if a player k leaves the game. In particular, consider all attacks y where $w_1(y) = k$ and $w_2(y) = i$. From $w_1(y)$'s definition, player i gains $\lambda_1^*(y)$ with k's exit (where $\lambda_1^*(y)$ denotes the new $\lambda_1(y)$ value with the imputed player set). The consequent difference in payoff $\Delta\Phi$ is

$$\Delta\Phi = \sum_{y \in \mathcal{Y}^t | w_1(y)=k} \log \lambda_1^*(y) - \log \lambda^t(k, i). \tag{29}$$

This corresponds to the payoff loss incurred by the leaving player k. Therefore, Φ reflects losses in a potential fashion.

Using the same rationale, it is easy to show that Φ also reflects losses correctly when i joins the game again. So Φ is a potential function.

In a potential game, the minima of the potential function Φ are Nash Equilibria. In this case, the equilibrium is the global minimum of Φ. Due to the potential function, the equilibrium can also be calculated iteratively with each new user entry, which makes the solution very suitable to study the evolution of user behavior with changing player counts.

References

Altman E, Boulogne T, El-Azouzi R, Jiménez T, Wynter L (2006) A survey on networking games in telecommunications. Comput Oper Res 33(2):286–311. https://doi.org/10.1016/j.cor.2004.06.005

Attias H (1999) A variational Bayesian framework for graphical models. In: Proceedings of the 12th international conference on neural information processing systems, NIPS'99. MIT Press, Cambridge, MA, USA, pp 209–215

Bala V, Goyal S (1998) Learning from neighbours. Rev Econ Stud 65(3):595–621. https://doi.org/10.1111/1467-937X.00059

Bar-Hillel A, Hertz T, Shental N, Weinshall D (2005). Learning a mahalanobis metric from equivalence constraints. J Mach Learn Res 6(32):937–965. URL http://jmlr.org/papers/v6/bar-hillel05a.html

Bellet A, Habrard A, Sebban M (2013) A survey on metric learning for feature vectors and structured data. CoRR, arXiv:abs/1306.6709

Benz A, Jäger G, Van Rooij R (2006) Game theory and pragmatics. Palgrave Macmillan, Basingstoke, New York, ISBN 9781403945723, 1403945721

Bishop CM (2006) Pattern recognition and machine learning. Springer, New York

Blackwell DA, Girshick MA (1980) Theory of games and statistical decisions. Dover Publications, New York

Borgs C, Chayes JT, Ding J, Lucier B (2011) The hitchhiker's guide to affiliation networks: a game-theoretic approach. In: Chazelle B (ed) Innovations in Computer Science - ICS 2011, Tsinghua University, Beijing, China, January 7-9. Proceedings, pp 389–400. Tsinghua University Press, 2011. http://conference.iiis.tsinghua.edu.cn/ICS2011/content/papers/22.html

Bosse T, Hoogendoorn M, Klein MCA, Treur J, van der Wal CN, van Wissen A (2013) Modelling collective decision making in groups and crowds: integrating social contagion and interacting emotions, beliefs and intentions. Autonom Agents Multi-Agent Syst 27(1):52–84. https://doi.org/10.1007/s10458-012-9201-1

Cha Youngchul, Cho Junghoo (2012) Social-network analysis using topic models. In *SIGIR'12*. ISBN 978-1-4503-1472-5. https://doi.org/10.1145/2348283.2348360

Chamley C (2004) Rational herds: economic models of social learning. Cambridge University Press, Cambridge ISBN 052182401X

Chen J, Wang G, Giannakis GB (2019) Nonlinear dimensionality reduction for discriminative analytics of multiple datasets. IEEE Trans Signal Process 67(3):740–752

Collobert R, Weston J, Bottou L, Karlen M, Kavukcuoglu K, Kuksa P (2011) Natural language processing (almost) from scratch. J Mach Learn Res 12. http://dl.acm.org/citation.cfm?id=1953048.2078186

Cosley D, Huttenlocher DP, Kleinberg JM, Lan X, Suri S (2010) Sequential influence models in social networks. In: Cohen WW, Gosling S (eds) Proceedings of the fourth international conference on weblogs and social media, ICWSM 2010, Washington, DC, USA, May 23–26, 2010. The AAAI Press. http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1530

Das S, Lavoie A (2014) The effects of feedback on human behavior in social media: an inverse reinforcement learning model. In: Bazzan ALC, Huhns MN, Lomuscio A, Scerri P (eds) International conference on Autonomous Agents and Multi-Agent Systems, AAMAS'14, Paris, France, May 5-9, 2014, pp 653–660. IFAAMAS/ACM. http://dl.acm.org/citation.cfm?id=2615837

Degroot MH (1974) Reaching a consensus. J Am Stat Assoc 69(345):118–121. https://doi.org/10.1080/01621459.1974.10480137

Demarzo P, Vayanos D, Zwiebel J (2003) Persuasion bias, social influence, and undimensional opinions. Q J Econ 118(3):909–968. ISSN 00335533. http://search.proquest.com/docview/211006209/

Dong H, Ping L, Zhong S, Liu C, Ji Y, Gong S (2018) Person re-identification by enhanced local maximal occurrence representation and generalized similarity metric learning. Neurocomputing 307:25–37. https://doi.org/10.1016/j.neucom.2018.04.013

Eliashberg J, Hui SK, Zhang ZJ (2014) Assessing box office performance using movie scripts: a kernel-based approach. IEEE Trans Knowl Data Eng 26(11):2639–2648

Franks H, Griffiths N, Anand SS (2014) Learning agent influence in mas with complex social networks. Autonom Agents Multi-Agent Syst 28(5):836–866. https://doi.org/10.1007/s10458-013-9241-1

Ghiassi M, Lio D, Moon B (2015) Pre-production forecasting of movie revenues with a dynamic artificial neural network. Expert Syst Appl 42(6):176–3193

Goldberger J, Roweis ST, Hinton GE, Salakhutdinov R (2004) Neighbourhood components analysis. In: Advances in neural information processing systems 17 [Neural Information Processing Systems, NIPS 2004, December 13-18, 2004, Vancouver, British Columbia, Canada], pp 513–520. https://proceedings.neurips.cc/paper/2004/hash/42fe880812925e520249e808937738d2-Abstract.html

Goodfellow I, McDaniel P, Papernot N (2018) Making machine learning robust against adversarial inputs. Commun ACM 61(7):56–66

Goyal A, Bonchi F, Lakshmanan LVS (2010) Learning influence probabilities in social networks. In: Davison BD, Suel T, Craswell N, Liu B (eds) Proceedings of the third international conference on web search and web data mining, WSDM 2010, New York, NY, USA, February 4-6, 2010. ACM, pp 241–250. https://doi.org/10.1145/1718487.1718518

Grandi U, Lorini E, Perrussel L (2015) Propositional opinion diffusion. In: Weiss G, Yolum P, Bordini RH, Elkind E (eds) Proceedings of the 2015 international conference on autonomous agents and multiagent systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015. ACM, pp 989–997. http://dl.acm.org/citation.cfm?id=2773278

Hillel AB, Weinshall D (2007) Learning distance function by coding similarity. In: Proceedings of the 24th international conference on machine learning, ICML'07, pp 65–72, New York, NY, USA. Association for Computing Machinery. ISBN 9781595937933. https://doi.org/10.1145/1273496.1273505

Hong L, Davison BD (2010) Empirical study of topic modeling in twitter. In: Giles CL, Mitra P, Perisic I, Yen J, Zhang H (eds) Proceedings of the 3rd workshop on social network mining and analysis, SNAKDD 2009, Paris, France, June 28, 2009. ACM, pp 80–88. https://doi.org/10.1145/1964858.1964870

Hu L, Wang X, Zhang M, Li J-Z, Li X, Shao C, Tang J, Liu Y (2015) Learning topic hierarchies for wikipedia categories. In: Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing of the Asian federation of natural language processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 2: Short Papers. The Association for Computer Linguistics, pp 346–351. https://doi.org/10.3115/v1/p15-2057

Korkmaz G, Kuhlman CJ, Marathe A, Marathe MV, Vega-Redondo F (2014) Collective action through common knowledge using a facebook model. In: AAMAS. ISBN 978-1-4503-2738-1. http://dl.acm.org/citation.cfm?id=2615731.2615774

Kostinger M, Hirzer M, Wohlhart P, Roth PM, Bischof H (2012) Large scale metric learning from equivalence constraints. In: 2012 IEEE conference on computer vision and pattern recognition, pp 2288–2295. IEEE. ISBN 9781467312264

Lash MT, Zhao K (2016) Early predictions of movie success: the who, what, and when of profitability. J Manag Inf Syst 33(3):874–903

Lerman K (2007) Social information processing in news aggregation. IEEE Internet Comput 11(6):16–28. https://doi.org/10.1109/MIC.2007.136

Milchtaich I (1996) Congestion games with player-specific payoff functions. Games Econ Behav 13(1):111–124. https://doi.org/10.1006/game.1996.0027

Mimno DM, McCallum A (2008) Topic models conditioned on arbitrary features with Dirichlet-multinomial regression. In: McAllester DA, Myllymäki P (eds) UAI. AUAI Press, pp 411–418. ISBN 0-9749039-4-9

Mit linguistics hyperlink constituency project. http://constituency.mit.edu/ (2014)

Moghaddam B, Jebara T, Pentland A (1998) Bayesian modeling of facial similarity. In: Kearns MJ, Solla SA, Cohn DA (eds) Advances in neural information processing systems 11, [NIPS Conference, Denver, Colorado, USA, November 30 - December 5, 1998]. The MIT Press, pp 910–916. http://papers.nips.cc/paper/1590-bayesian-modeling-of-facial-similarity

Nisan N, Roughgarden T, Tardos E, Vazirani VV (2007) Algorithmic game theory. Cambridge University Press, New York (**ISBN 0521872820**)

O'Hagan A, Forster JJ (2004) Kendall's advanced theory of statistics, volume 2B: Bayesian Inference, 2nd edn, vol 2B

Olfati-Saber, Alex Fax J, Murray RM (2007) Consensus and cooperation in networked multi-agent systems. Proceedings of the IEEE 95(1):215–233. https://doi.org/10.1109/JPROC.2006.887293

Panait L, Luke S (2005) Cooperative multi-agent learning: the state of the art. Autonom Agents Multi-Agent Syst 11(3):387–434. https://doi.org/10.1007/s10458-005-2631-2

Pennacchiotti M, Gurumurthy S (2011) Investigating topic models for social media user recommendation. In: Srinivasan S, Ramamritham K, Kumar A, Ravindra MP, Bertino E, Kumar R (eds) Proceedings of the 20th international conference on world wide web, WWW 2011, Hyderabad, India, March 28–April 1, 2011 (Companion Volume). ACM, pp 101–102. https://doi.org/10.1145/1963192.1963244

Radinsky K, Svore KM, Dumais ST, Teevan J, Bocharov A, Horvitz E (2012) Modeling and predicting behavioral dynamics on the web. In: Mille A, Gandon F, Misselis J, Rabinovich M, Staab S (eds) Proceedings of the 21st world wide web conference 2012, WWW 2012, Lyon, France, April 16-20, 2012. ACM, pp 599–608. https://doi.org/10.1145/2187836.2187918

Rosen-Zvi M, Griffiths T, Steyvers M, Smyth P (2004) The author-topic model for authors and documents. In: Proceedings of the 20th conference on uncertainty in artificial intelligence, UAI '04, pp 487–494, Arlington, Virginia, United States. AUAI Press. ISBN 0-9749039-0-6. http://dl.acm.org/citation.cfm?id=1036843.1036902

Segaran T (2007) Programming collective intelligence: building smart web 2.0 applications. O'Reilly, Sebastapol, CA. ISBN 1-306-81760-9; 0-596-51760-2; 0-596-55068-5

Singh S, Subramanya A, Pereira F, McCallum A (2012) Wikilinks: a large-scale cross-document coreference corpus labeled via links to Wikipedia. Technical Report UM-CS-2012-015

Steels L (2012) Experiments in cultural language evolution, vol 3. John Benjamins Publishing Company, Philadelphia. ISBN 9789027204561; 902720456X

Steinberg DM, Pizarro O, Williams SB (2015) Hierarchical Bayesian models for unsupervised scene understanding. Comput Vis Image Underst 131(C):128–144. https://doi.org/10.1016/j.cviu.2014.06.004

Szabó G, Huberman BA (2010) Predicting the popularity of online content. Commun ACM 53(8):80–88. https://doi.org/10.1145/1787234.1787254

Tsang A, Larson K (2014) Opinion dynamics of skeptical agents. In: Bazzan ALC, Huhns MN, Lomuscio A, Scerri P (eds) International conference on autonomous agents and multi-agent systems, AAMAS'14, Paris, France, May 5-9, 2014, pp 277–284. IFAAMAS/ACM. http://dl.acm.org/citation.cfm?id=2615778

Wang J, Gasser L (2002) Mutual online concept learning for multiple agents. In: The first international joint conference on autonomous agents & multiagent systems, AAMAS 2002, July 15-19, 2002, Bologna, Italy, Proceedings. ACM, pp 362–369. https://doi.org/10.1145/544741.544830

Weinberger KQ, Saul LK (2009) Distance metric learning for large margin nearest neighbor classification. J Mach Learn Res 10. ISSN 1532-4435. http://dl.acm.org/citation.cfm?id=1577069.1577078

Wikimedia (2012) Wikimedia page view counts. https://dumps.wikimedia.org/other/pagecounts-raw/

Wikipedia participation challenge (2012) http://www.kaggle.com/c/wikichallenge

Xtag morphology database (2014) http://www.cis.upenn.edu/~xtag/

Zhang H, Zhang S, Wu Z, Huang L, Ma Y (2014) Predicting wikipedia editor's editing interest based on factor graph model. In: IEEE international congress on big data. ISBN 978-1-4799-5057-7. https://doi.org/10.1109/BigData.Congress.2014.63

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.