SAND2017-6662C

Estimating users' mode transition functions and activity levels from social media

Hamilton Link, Jeremy D. Wendt, Richard V. Field, Jr., and Jocelyn Marthe
Sandia National Laboratories
Albuquerque, New Mexico U.S.A.
Email: {helink, jdwendt, rvfield, jmmarthe}@sandia.gov

***Abstract*—We present a temporal model of individual-scale social media user behavior, comprising modal activity levels and mode switching patterns. We show that this model can be effectively and easily learned from available social media data, and that our model is sufficiently flexible to capture diverse users' daily activity patterns. In applications such as electric power load prediction, computer network traffic analysis, disease spread modeling, and disease outbreak forecasting, it is useful to have a model of individual-scale patterns of human behavior. Our user model is intended to be suitable for integration into such population models, for future applications of prediction, change detection, or agent-based simulation.**

I. Introduction

On-line social media contain massive amounts of latent information of benefit to building user models. When trained and used to track the conditional likelihood of new content, user models can be applied in many ways. Herein, we present a Bayesian model that interprets timestamped events in the context of a hidden user mode. Our work focuses on representing daily shifts in per-user temporal patterns. A hidden Markov model tracks the user's changing mode, which affects the interval between events; we add to this conventional baseline a learned function for the user's shifting daily mode transition probabilities.

A Bayesian network represents the dependencies between hidden domain structure and observable data. Once learned, networks may be used in forecasting and active analysis. As an example, when using social media for disease forecasting, we might learn activity probabilities from user data and then simulate varying disease transmission on a society scale. When building models for change detection, we might use similar models to monitor sensor data likelihood under several competing scenarios of interest. At the heart of Bayesian analysis are algorithms for approximation of the posterior hidden-value likelihood distribution supported by the evidence. In this work we used a Markov chain Monte Carlo (MCMC) strategy known as "Gibbs sampling" [2] both as a parameter approximation algorithm and to generate statistics for comparison with training data as a posterior predictive check of the model.

We use Twitter as the source of our per-user temporal streams. Twitter is a popular microblogging site where users regularly generate *tweets* (very brief and generally public status updates) that are immediately posted for others to read. As these tweets require minimal effort to generate and are posted immediately, we believe their temporal patterns can be useful for observing changes and patterns in a society's day-to-day activities. Twitter provides a public API which allows access to large volumes of tweets [4] to enable social network research and development of media-enabled services.

We present a model for an individual user's temporal behavior (Section III) and show that it accurately fits real users' temporal patterns (Section IV). This model is relatively straightforward, directly implementable in modern statistical languages, and runs efficiently when fitting model parameters to data.

II. Previous Work

Many have studied Twitter for its own sake, or as a proxy for phenomena in real-world (vs. on-line) social networks and human communication. For example, Lukasik *et al.* studied tweet arrival times across users on the same topic/rumor using inhomogenous Poisson processes [5]. Their model and our own are Bayesian and use log-normal distributions but model different phenomena. Additional types of stochastic process models have been used to study information propagation through social media [6]–[8], including cases where information may pass out of band (through other social networks, human interaction, advertising, *etc.*). Yang and Zha [9] present an example of this using variational Bayes to simultaneously track spreading memes while estimating the diffusion network through which the meme was spreading.

Our work was inspired by Watson's Time Maps [10] – a visualization of a single user's temporal Twitter usage patterns. Figure 1 shows a Time Map for a single user. Others have used this as a visualization technique to study users (*e.g.*, [11]) but not as the basis for graphical time series models.

In some ways, our work could be compared with Linderman et al.'s recurrent switching linear dynamical systems [12]. They split temporal data into modes based on current mode and other covariates. However, their work includes linear dynamic systems (*e.g.*, basketball player trajectories) while ours focuses purely on temporal dynamics (*e.g.*, when one may be more likely to tweet more often) and our model is therefore simpler.

There are a number of strategies for forecasting the severity of outbreaks at a range of scales. Examples include aggregated diffusion models informed by airline traffic networks

[13], regional movement models informed by cell tower data [14], and detailed agent-based models representing millions of individuals incorporating subject matter expertise [3]. In EpiSimS, each individual was assigned a schedule comprised of eight activities, and each individual's propensity to isolate themselves when ill was prescribed. Bayesian sampling algorithms are typically computationally demanding, so for massive simulation applications such as [3] it would be more appropriate to sample derivative quantities from the model and produce pregenerated schedules than produce random events directly from the posterior distribution.

For many diseases, managing them is greatly aided by early detection. Social media, phone centers, and internet search records have been used to provide early warning indications of outbreaks when people begin to seek out information or help [1], [15]. Our proposed model could be incorporated into a system that detects subtle behavioral shifts and could potentially also contribute to such early warning systems for disaster response [16].

III. Proposed Model

We present a model that accurately represents much of Twitter users' temporal behaviors while being simple enough to use in agent-based simulations. This model builds upon the assumption that each user has a discrete set of activity levels, or temporal "modes", within which the time between tweets is predictably distributed. The parameters governing these activity levels are estimated from the data along with coefficients that describe the user's likelihood of switching between modes based on the time of day.

A simple example will help illustrate our point. Over the course of a normal day, user Alice tweets every few hours. Occasionally, however, user Bob invites Alice to join him and a group of friends to dinner via a targeted message (known as an "at"-refer or @-refer). Alice responds to Bob's message, and during the exchange they send several tweets within minutes of each other discussing alternatives, and confirming the time, location, and other details. At the end of most days, Alice ceases tweeting in the late evening, and is not likely to resume tweeting until the following morning.

In this example, Alice exhibits three modes of activity – "engaged" (tweets every few minutes), "casual" (tweets every few hours), and "sleeping" (many hours until next tweet). Moreover, Alice's likelihood to enter the "sleeping" mode increases throughout the day, and while she is very likely to leave this mode in the morning, it is left to the model to estimate her predisposition to begin the day by engaging in new discussions or tweeting casually. Although it may seem contrived, patterns such as the one described can be found in many users' post time data on Twitter. To illustrate, Fig. 1 charts the time between tweets for a single individual whom we refer to as User 28; the distribution of tweets match this description particularly well. In practice, some users may exhibit more or less than three modes. To avoid reading too much into the results of any individual user, we will simply

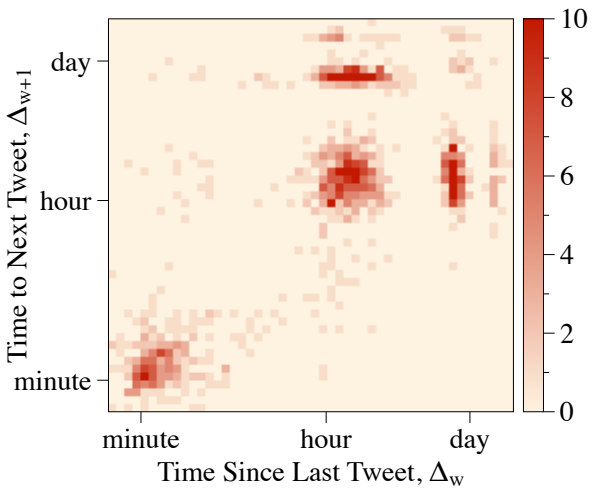

Fig. 1. **Time map of pronounced modes.** This user (User 28) has three (uncharacteristically) distinct and clear temporal modes: "engaged" in the lower left corner and "casual" near the center of the plot, with a faint "sleeping" mode in the upper right corner. The two hot spots above and to the right of the central casual mode indicate that User 28 generally enters the "sleeping" mode from the casual mode, and typically returns to the casual mode. Note that both axes are at logarithmic scale.

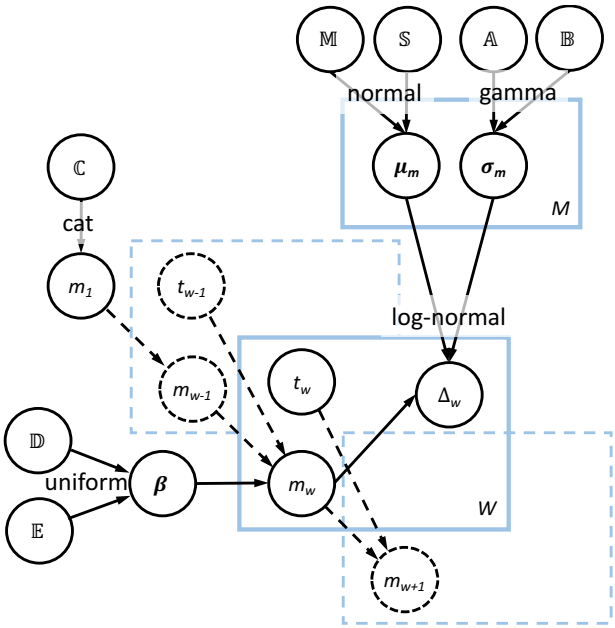

Fig. 2. **Plate notation for single-user model.** μ, σ, m, and β are hidden variables. t and Δ are observed. When expanding this shorthand description, nodes in plates M and W are replicated and indexed based on the size of the corresponding sets, and selectively reconnected based on the conditional dependencies described. [17]

TABLE I
NOTATION DEFINITIONS.

W	Set of user's tweets
$\|W\|$	Number of user's tweets
w	Tweet index, $w \in 1 \dots \|W\|$
t_w	Time of day of tweet w
Δ_w	Previous interval $= t_w - t_{w-1}$
M	Set of user's modes
$\|M\|$	Number of user's modes
m_w	Mode ordinal of tweet w
μ_m	Mean of mode m
σ_m	Standard deviation of mode m
β	Transition function coefficients
$\mathbb{M}, \mathbb{S}$	Prior parameters for μ
$\mathbb{A}, \mathbb{B}$	Prior parameters for σ
$\mathbb{C}$	Categorial parameter for m_0
$\mathbb{D}, \mathbb{E}$	Prior bounds on elements of β

refer to the modes by their ordinal numbers; modes are sorted in increasing mean time intervals to avoid aliasing.

We use Bayesian methods to model the different user modes and transition likelihoods [18], [19]. Figure 2 shows our proposed model in plate notation, and Table I provides the meaning of each of the variables. A set of tweets (plate W, indexed with w) are observed, each at some time of day (t_w, minutes into the day), generated by some unknown activities we describe as a set of modes (plate M, indexed with m, and corresponding to the possible values of each random variable m_w). The mode m_w producing a tweet is selected based on the time of the previous tweet, the previous mode m_{w-1}, and a function that models the shifting probabilities of mode transitions from m_{w-1} to m_w. This hidden random function is a categorical-valued function of time of day and mode, represented with hidden variable coefficients β. Mode m_w determines the distribution for the time interval between this tweet and the previous, denoted by Δ_w, which is observed for all but the first tweet m_1. Under this model, to make clear a possible point of confusion, the time of day and mode of each tweet (t_{w-1} and m_{w-1}) along with the learned function coefficients β determine the relative probabilities of the *next* mode m_w; so, a user's casual tweet late in the evening may make it more likely for the next mode to be "sleeping." However, that mode then determines the Δ_w delay *leading up to* the associated tweet; so, a user's first message of the morning might be produced by the "sleeping" mode making a lengthy interval that spanned the night more likely.

The posterior probability of all model parameters is given by

$$
\begin{aligned}
p(\Delta_*, m_*, \mu_*, \sigma_*, \beta | t_*, \mathbb{M}, \mathbb{S}, \mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}, \mathbb{E}) \propto \\
&\prod_{w=2}^{|W|} \mathcal{N}(\log(\Delta_w) | \mu_{m_w}, \sigma_{m_w}) \\
&\times \prod_{w=2}^{|W|} p(m_w | m_{w-1}, t_{w-1}, \beta) \\
&\times \prod_{m=1}^{|M|} \left[\mathcal{N}(\mu_m | \mathbb{M}, \mathbb{S}) \, \text{Gamma}(\sigma_m | \mathbb{A}, \mathbb{B}) \right] \\
&\times \text{Cat}(m_1 | \mathbb{C}) \, p(\beta | \mathbb{D}, \mathbb{E}),
\end{aligned}
\tag{1}
$$

where $\mathcal{N}$, Gamma, and Cat denote the probability density functions for the standard normal, gamma, and categorial distributions, respectively. The asterisk subscript (*e.g.*, Δ_*) indicates use of the elements of the vector in question. The Δ_* are observed for $w \neq 1$ and the covariates t_* are fully observed. These are separated only for convenience, the former being entirely determined by the latter.

The mode transition factors $p(\beta | \mathbb{D}, \mathbb{E})$ and $p(m_w | m_{w-1}, t_{w-1}, \beta)$ in Eq. (1) incorporate a variable-

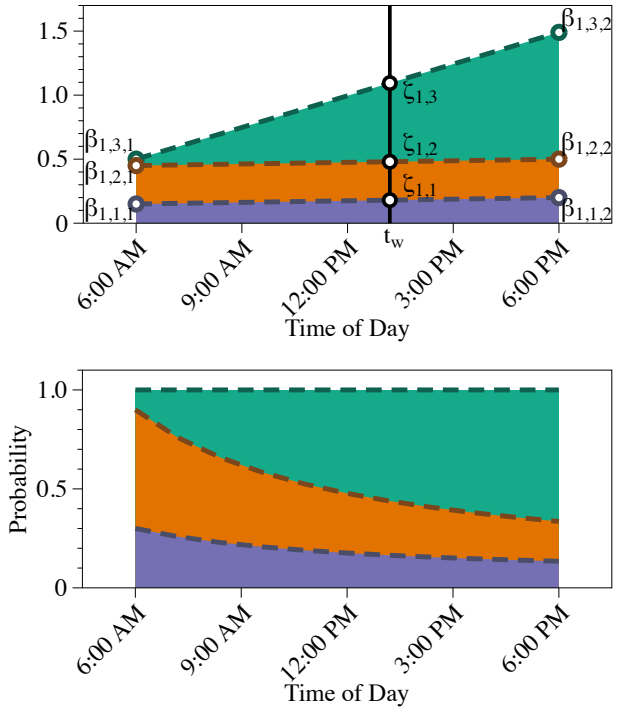

Fig. 3. **Mode transition probability at time t_w for a single mode $m_w = 1$.** An example of how mode transition probability is computed from a set of coefficients $\beta_{m_{w-1}, m_w, i}$. See Eqs. (2a)–(2d).

valued function of time of day broken down as:

$$p(m_w|m_{w-1}, t_{w-1}, \beta) \sim \text{Cat}(\zeta_{r=(m_{w-1},*)}) \quad (2a)$$

$$\zeta_r = \text{lerp}(\beta_{r,i},\ \beta_{r,j},\ t_w) \quad (2b)$$

$$p(\beta|\mathbb{D}, \mathbb{E}) = \prod_{r,i} p(\beta_{r,i}|\mathbb{D}, \mathbb{E}) \quad (2c)$$

$$p(\beta_{r,*}|\mathbb{D}, \mathbb{E}) \sim \text{U}(\mathbb{D}, \mathbb{E}), \quad (2d)$$

where U denotes the probability density function for the uniform distribution, $r = (m_{w-1}, m_w)$ is a notational convenience to designate the appropriate mode pair, and i (or j) is a lower (or upper) bound on the time of day messages are posted. In Eq. (2a), $\zeta_{r=(m_{w-1},*)}$ is a vector of calculated values given the previous mode and each possible subsequent mode. In Eq. (2b), $\text{lerp}(a, b, c)$ is the linear interpolation function between a and b at a distance c. More precisely:

$$\text{lerp}(\beta_{r,i},\ \beta_{r,j},\ t_w) = \beta_{r,i} + (t_w - i)\frac{\beta_{r,j} - \beta_{r,i}}{j - i} \quad (3)$$

Figure 3 depicts how categorical mode probabilities are drawn from these equations. During sampling, the generated coefficients β implicitly define ratios which, when normalized, allow for a certain amount of nonlinear behavior in the shifting mode transition probabilities throughout the day. The specific β-derived ratio for the time of the current tweet is used for likelihood of transitioning into any next mode (Eqs. (2a) and (2b)).

Having defined the posterior likelihood function over the entire model, the next step was to implement the model in JAGS [2]. The implementation tells JAGS how to calculate the total data log likelihood given candidate values of all hidden variables. The performance of the model was sufficient to render such mathematical exercises unnecessary.

IV. Data and Results

Using the Twitter API, we selected 29 users that included a mix of organizations, celebrities, automated systems ("bots"), and conventional individual users,[1] as well as the time stamps of their most recent tweets.[2] For privacy concerns, we do not provide identifiers and instead refer to all of these entities as "users". We sought out several different types of users to test the model's flexibility; collecting representative samples that would enable good population distribution estimates is beyond the scope of this paper.

We implemented the model defined by Eq. (1) using R [20] and JAGS (via rjags [2]) to describe the likelihood calculations. JAGS develops a Gibbs sampler, which produces random samples from a converging estimate of the posterior joint probability distribution across all hidden variables. Because tweets are not posted with a time zone or the sender's local time, the different users live in different time zones and have varying sleep patterns, a procedure to align the data is needed. We computed each user's time-of-day tweeting histogram by shifting their associated timezone such that midnight in the shifted data corresponded with the time period they were least likely to tweet. In a larger study it would be desirable to extend the model to estimate the time zone of each user or alter the model transition likelihood function in Eq. (2b) to be periodic (or both) and avoid this extra step. In simulation, individual agent time zones could be resampled from a distribution or initialized deterministically depending on the application.

For each user, we varied the number of modes as $|M| \in \{2, 3, 4\}$ for a total of 87 results, running four chains each with 1000 samples for warm-up and 1000 samples for the fit, for a total of 4000 samples for each hidden variable in the posterior joint probability distribution. The predetermined number of modes $|M|$ used to fit the data had a significant impact on the run time. Average runtimes were 6, 11, and 17 minutes for increasing $|M|$. Once complete, we studied all results and manually assigned each user a value for $|M|$ that best fit their data.

Table II lists the model parameters, chosen to produce uninformed or weakly informed priors. The value of parameters $\mathbb{D}$ and $\mathbb{E}$ are not significant, and the sampler appeared to find reasonable posterior ratios which describe the mode transition probabilities (as shown, for example, in Fig. 5). In the case of $\mathbb{M}$, the prior mode means, the four values considered correspond approximately to one minute, one hour, two days, and 100 days. We set the log-normal standard deviations $\mathbb{S}$ on the prior larger than the difference between the values $\mathbb{M}$. We will discuss the effect of $\mathbb{M}$ on the results in more detail

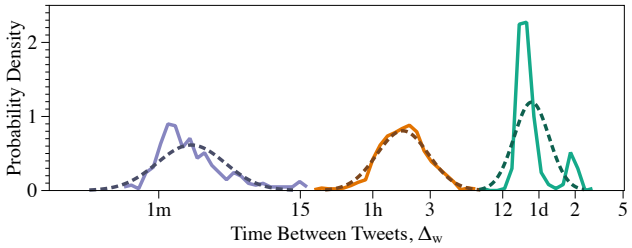

Fig. 4. **Three-mode model of User 28.** In this case the data is clearly separable into the three modes (color of raw histograms) and a mixture of normal distributions is generally a good fit to the raw data, although mode 3 is used to span both one-evening and two-day intervals.

TABLE II
Model Parameters.

$\mathbb{M}$	$\{4n : n \in 1...\|M\|\}$
$\mathbb{S}$	5
$\mathbb{A}$	1
$\mathbb{B}$	1
$\mathbb{C}$	$1/\|M\|$
$\mathbb{D}$	0.001
$\mathbb{E}$	0.999

[1]Users other than organizations and celebrity personalities were chosen by selecting a random ID from recent tweets published by the Twitter drip line, which is a non-uniform random sample from 1% of all tweets. This technique is biased toward users who tweet more often, and returns both apparent real users and apparent "bots".

[2]Generally 1,000–1,200 tweets per user ID; two users had tweeted fewer times (245, 690).

below.

In Figs. 4 and 5, we further study the tweet pattern of User 28. By Fig. 4, a log-normal distribution provides a good fit to the data on time between tweets. Specifically, we plot the sample median (50%) values for μ_m and σ_m (we provide bounds for these fits in later plots). The data is plotted as probability density functions with the data split and colored based on the most likely mode given the intervals defined by μ_m and σ_m. While User 28 has three clear and separable modes, not all users do. In Fig. 6, we illustrate the tweet pattern of User 13, whose second mode intersects the first and third modes. User 21, illustrated by Fig. 7, exhibits two modes which intersect.

Figure 5 illustrates the mode-transition fits for User 28. The horizontal axis is split into $|M|$ pieces with each piece showing transitions *from* the specified mode into all modes. The background bars represent the original data binned throughout the day with stacked bars showing which portion of each bin transitions into all modes. The black dashed lines show the 50% fit for the β function dividing modes. That is, below the first line are the transitions into mode 1; between the two lines are transitions into mode 2; above both lines are transitions into mode 3. We also show the confidence interval from 5% to 95% on each of those fits. The top portion of the plot shows the histogram of how many datapoints are included in each of the data bins below.[3] The model is biased to fit portions of the day where more data exists. We note that Figs. 6 and 7 are similar and have this same feature; the confidence intervals were removed from these figures for clarity.

In Fig. 5, we see that User 28 has a distinct trend toward leaving mode 3 and entering mode 2 first thing in the morning. Although User 28 has a low probability of leaving mode 2 to enter mode 1 at any point in the day, once in mode 1, this user is very likely to remain there. In all three modes, the likelihood of entering mode 3 increases throughout the day.[4]

Figures 8–10 show values for μ_m and σ_m for all users. For each user, we determined the optimal number of modes $|M|$ and organized the figures accordingly. In all plots, the boxes show the 5%, 50%, and 95% likelihoods for the μ_m values; the whiskers show the 50% value for σ_m centered on the 50% μ_m. The vertical dotted lines indicate locations for the $\mathbb{M}$ prior value for all μ_m. Note that with a prior $\mathbb{S}$ of 5 (Table II), the μ_m values should be rather free to follow the data away from the prior $\mathbb{M}$. We note that the model often moved quite far from the priors. Furthermore, with the exceptions of Users 1, 8, 22, and 25, all modes' μ_m 90% confidence intervals were very tight. In around half of the cases, the modes' σ_m values are also quite separable – there are often clear divisions between the separate modes.

The results for User 28 are particularly interesting. This is the only user to be in more than one of these plots (Figs. 9 and 10). User 28 is also shown in Fig. 4 – where it can be seen that a possible fourth mode is visible. We show the results of the 4-mode fit in Fig. 10. The fourth mode is pulled toward the fourth mode's data, but there was insufficient data for the model to completely shift over. Moreover, the standard deviation on the fourth mode is exceedingly high. Only two users were reasonably modeled using a 4-mode model – and in User 28's case, the fourth mode was insufficiently strong to lead to a confident fit. This is a standard limitation of temporal likelihood models; longer intervals produce fewer samples upon which to base the likelihood estimate of new data for any application. Even with good population estimates of rare phenomena, individual deviations from the norm in social media will remain difficult to evaluate.

We observed that for some users who clearly displayed two modes, the model did not distinguish the modes as well as we might hope (Fig. 11). Our initial intent with our prior parameters was that weakly informed priors would easily be shifted by sufficient data. However, these results indicate that some users' data was insufficient to fully overcome the priors. As a test, we took these users and tested them with slightly different priors ($\mathbb{M} = \{4, 12\}$ instead of $\mathbb{M} = \{4, 8\}$). In all cases the resulting fits were improved, and in most cases, the splits were now as expected. Further, the second mode was far less common – resulting in far less data – than the first mode. We show the resulting modes for all of these users in Fig. 12.

Finally, "bots" and otherwise automated accounts can sometimes cause problems for our model. Figure 13 illustrates the tweet patterns for Users 18 and 19. The top plot illustrates User 18, who nearly always tweets every few seconds, and tweets every 2-3 hours much less frequently. As the modes were very disjoint, the model was able to identify the two modes well. However, the User 19 (bottom panel) tweets at very precise intervals (approximately 5 minutes, 10 minutes, 15 minutes, and 30 minutes) with far fewer tweets to fill in around them. In the two-mode case, neither of the modes was a good fit to the data. When considering more than two modes, the fit for at least one of the modes exhibited a variance near zero (essentially a Dirac delta function) while others were unable to converge. We assume the near zero value for σ_m led to numerical issues for the solver.

V. Conclusions

We presented a new Bayesian model that effectively fits shifting temporal patterns as seen in Twitter. This model integrates a user's different temporal modes ("engaged", "casual", "sleeping") with time-of-day interactions. Moreover, the model is relatively straightforward and can be easily implemented.

We demonstrated how well the proposed model fit a variety of users by analyzing the tweets of 29 different users with several model variations; results for a subset of these users were presented. For some user modes, there was insufficient data to overwhelm the priors provided. Low energy (lower frequency) modes will necessarily produce fewer samples; this inherant lack of data may make other distributions (such as Student-t) more appropriate for estimating the parameters of such modes,

[3]While our plots compare posterior estimates to manually binned data, we should make clear the Gibbs sampler included randomly resampling individual message modes and did not use our labels.

[4]This was a fairly common trend among most non-"bot" users: As the day progressed, they were more likely to enter their longest between-tweet mode.

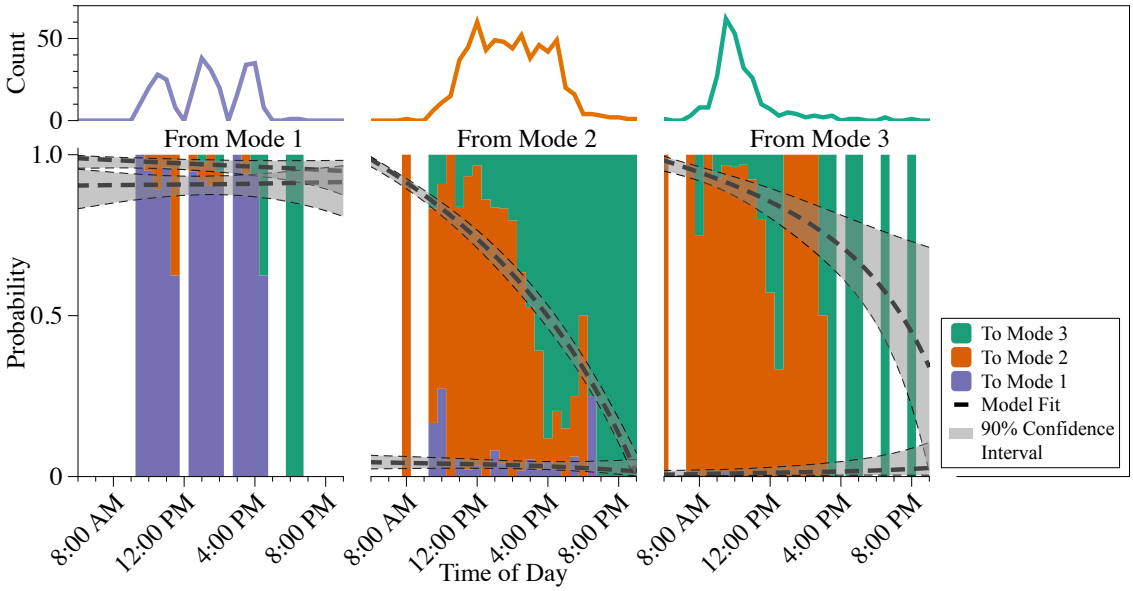

Fig. 5. **Transition fits with data.** A 3-mode fit to the mode transitions for User 28. In the top panel are histograms for each mode, of the number of tweets generated in that mode by time of day. In the bottom panel, those same tweets are subdivided further in a stacked normalized bar chart showing the fraction of occurrences of each possible next mode. From left to right each pair of plots illustrate transitions from one of the three modes. Thus, the purple bars in the left plot are those tweets that transitioned from mode 1 to mode 1. The black ribbons show the model's estimated transition probability functions including uncertainty. Note the bar charts in these plots are based on our *assigned* labels for the data; the estimated functions are based on *sampled* labels for this data.

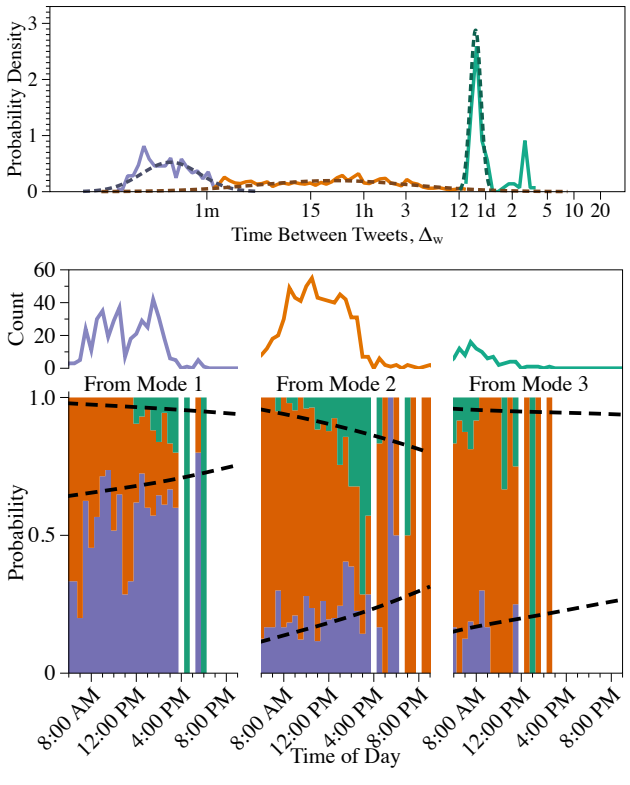

Fig. 6. **Three-mode model of User 13.** Note that this user's behavior is similar to User 28 shown in previous figures, but with less regularity and separability between modes.

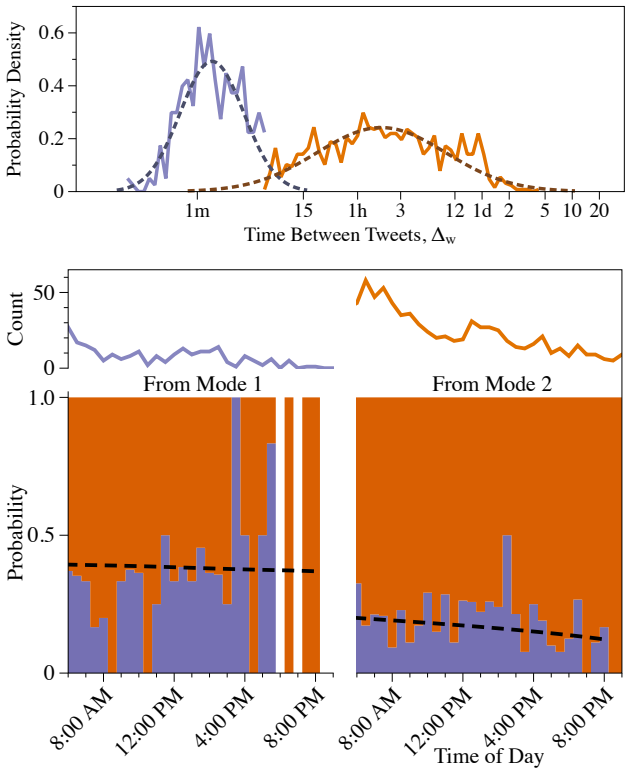

Fig. 7. **Two-mode model of User 21.**

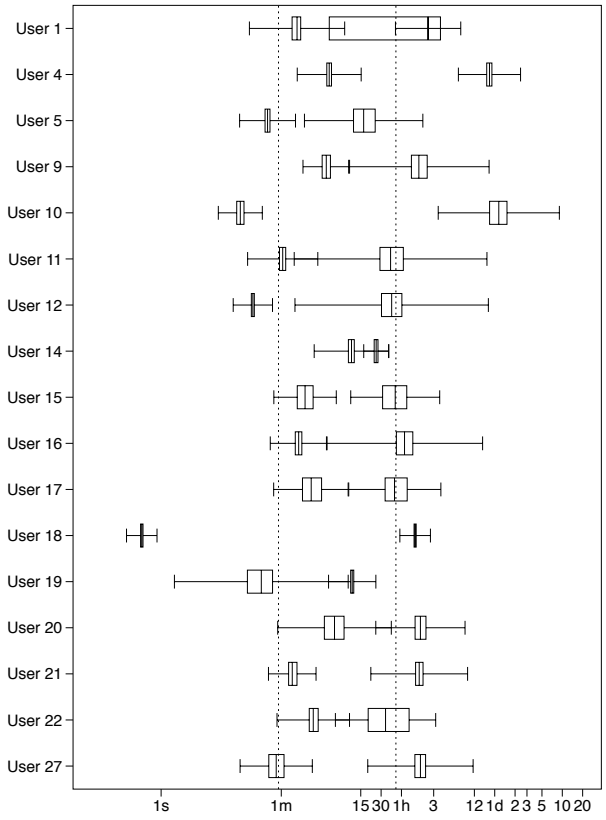

Fig. 8. **Two-mode user models.** This shows all users who were best fit by 2 modes. The boxes show the 5%, 50%, and 95% confidence intervals on μ_m for each mode. The whiskers show the 50% estimates for σ_m as offset from the 50% estimate for μ_m. The vertical dotted lines show the hyperparameter $\mathbb{M}$ for each μ_m.

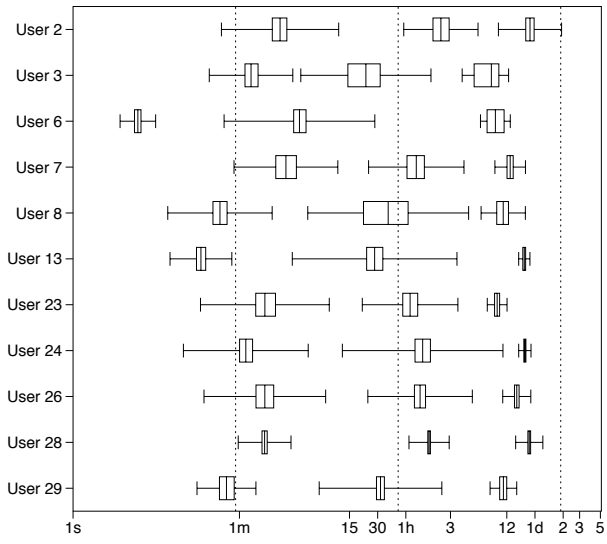

Fig. 9. **Three-mode user models.** This shows all users who were best fit by 3 modes. The boxes show the 5%, 50%, and 95% confidence intervals on μ_m for each mode. The whiskers show the 50% estimates for σ_m as offset from the 50% estimate for μ_m. The vertical dotted lines show the hyperparameter $\mathbb{M}$ for each μ_m.

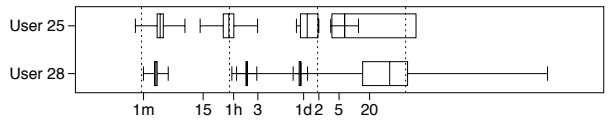

Fig. 10. **Four-mode user models.** This shows all users who were best fit by 4 modes. The boxes show the 5%, 50%, and 95% confidence intervals on μ_m for each mode. The whiskers show the 50% estimates for σ_m as offset from the 50% estimate for μ_m. The vertical dotted lines show the hyperparameter $\mathbb{M}$ for each μ_m.

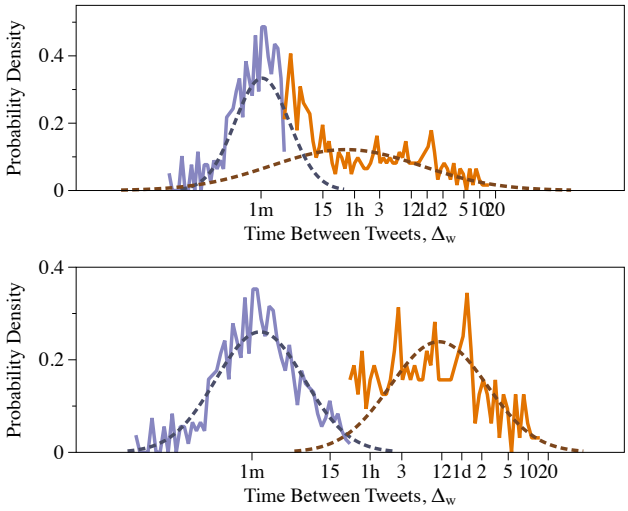

Fig. 11. **Two-mode model sensitivity to priors.** The two modes evident in User 11's tweet pattern were not well split by our default priors (top panel), but after shifting the priors the model produced a more satisfactory result (bottom panel).

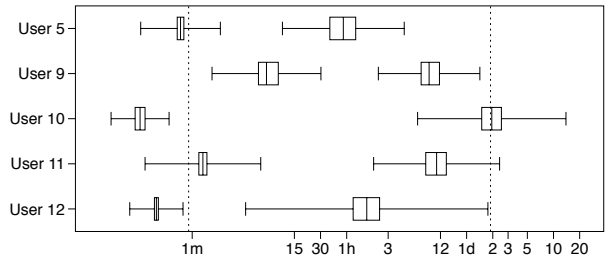

Fig. 12. **Two-mode models, shifted priors.** This shows the resulting fits with slightly altered priors for a small number of users whose fits were not well served by our default priors. The boxes show the 5%, 50%, and 95% confidence intervals on μ_m for each mode. The whiskers show the 50% estimates for σ_m as offset from the 50% estimate for μ_m. The vertical dotted lines show the hyperparameter $\mathbb{M}$ for each μ_m.

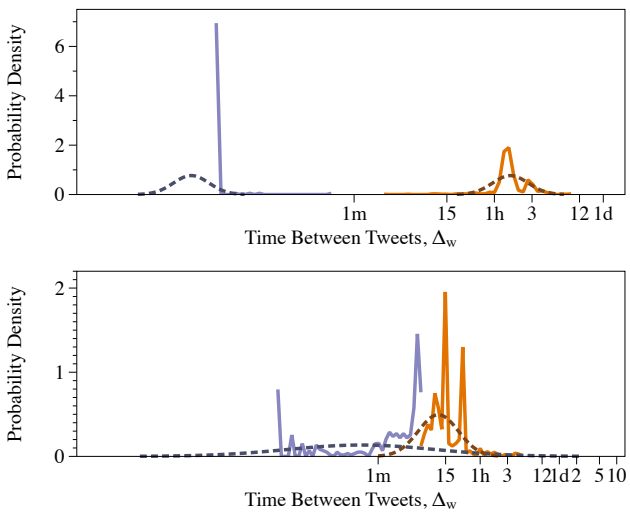

Fig. 13. **Resulting fits to two automated accounts.** The tweeting behavior for User 18 (top panel) and User 19 (bottom panel) suggest significant automated behavior. In the top panel, the two modes are sufficiently separated that the model still provides a reasonable fit. In the bottom, the modes are not well represented by a log-normal mixture model, and no good fit could be found.

and may make uninformative priors preferable to even weakly informative priors (absent concrete prior knowledge). Zero- or near-zero variance modes (commonly occurring in posts by automated systems) may be better modeled with improper priors and delta function or custom likelihood calculations, to avoid misleading compromises in posterior variance estimates. We believe further study is required to determine how choice of priors and variable distributions affects these results.

While weaker priors could allow less data to lead to good results for some users, they may cause degradation of some of our good results. An obvious model extension that doubles as a potential solution to this problem is to explicitly represent structural variations between users within an overarching population, sampling user-scale parameters from hidden population distributions. Such a model could go so far as to automatically identify user classes by clustering users' temporal patterns in parallel with user parameter estimation and population parameter estimation.

Based on the results presented here, we are confident that a successful model should be able to reliably represent organizations, "bots", and human users, and potentially identify subgroups within these broad demographics. It should also be useful for pregenerating schedules for agent-based simulations or change detection in early warning applications.

Acknowledgments

This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

References

[1] T. Ahmad, N. A. Rehman, F. Pervaiz, S. Kalyanaraman, M. B. Safeer, S. Chakraborty, U. Saif, and L. Subramanian, "Characterizing dengue spread and severity using internet media sources," in *Proceedings of the 3rd ACM Symposium on Computing for Development*. Bangalore, India: ACM, January 2013.

[2] M. Plummer. (2016) Package RJAGS. [Online]. Available: http://cran.r-project.org/web/packages/rjags/rjags.pdf

[3] S. M. Mniszewski, S. Y. Del Valle, P. D. Stroud, J. M. Riese, and S. J. Sydoriak, "EpiSimS simulation of a multi-component strategy for pandemic influenza," in *Proceedings of the 2008 Spring Simulation Multiconference*, ser. SpringSim '08. San Diego, CA, USA: Society for Computer Simulation International, 2008, pp. 556–563. [Online]. Available: http://dl.acm.org/citation.cfm?id=1400549.1400636

[4] Twitter. (2017, January). [Online]. Available: https://dev.twitter.com/docs/

[5] M. Lukasik, P. Srijith, T. Cohn, and K. Bontcheva, "Modeling tweet arrival times using log-gaussian Cox processes." in *Proceedings of Emperical Methods of Natural Language Processing*, 2015, pp. 250–255.

[6] Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, "SEISMIC: A self-exciting point process model for predicting tweet popularity," in *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1513–1522.

[7] J. Gao, H. Shen, S. Liu, and X. Cheng, "Modeling and predicting retweeting dynamics via a mixture process," in *WWW'16 Companion*. Montréal, Québec, Canada: ACM, April 2016.

[8] A. Simma and M. I. Jordan, "Modeling events with cascades of poisson processes," *arXiv preprint arXiv:1203.3516*, 2012.

[9] S. Yang and H. Zha, "Mixture of mutually exciting processes for viral diffusion," in *Proceedings of the 30th International Conference on Machine Learning*, ser. JMLR W&CP, vol. 28, Atlanta, GA, 2013.

[10] M. C. Watson, "Time maps: A tool for visualizing many discrete events across multiple timescales," in *Proceedings of the 2015 IEEE International Conference on Big Data*. IEEE, 2015, pp. 793–800.

[11] N. M. Radziwill and M. C. Benton, "Bot or not? deciphering time maps for tweet interarrivals," arXiv:1605.06555 [cs.SI], 2016.

[12] S. W. Linderman, A. C. Miller, R. P. Adams, D. M. Blei, L. Paninski, and M. J. Johnson, "Recurrent switching linear dynamical systems," arxiv:1610.08466, 2016.

[13] K. B. Gustafson, B. S. Bayati, and P. A. Eckhoff, "Fractional diffusion emulates a human mobility network during a simulated disease outbreak," *arXiv preprint arXiv:1601.07655*, 2016.

[14] A. Wesolowski, N. Eagle, A. J. Tatem, D. L. Smith, A. M. Noor, R. W. Snow, and C. O. Buckee, "Quantifying the impact of human mobility on malaria," *Science*, vol. 338, no. 6104, pp. 267–270, 2012.

[15] N. A. Rehman, S. Kalyanaraman, T. Ahmad, F. Pervaiz, U. Saif, and L. Subramanian, "Fine-grained dengue forecasting using telephone triage services," *Science Advances*, vol. 2, no. 7, 2016.

[16] P. S. Earle, D. C. Bowden, and M. Guy, "Twitter earthquake detection: earthquake monitoring in a social world," *Annals of Geophysics*, vol. 54, no. 6, pp. 708–715, October 2011.

[17] W. L. Buntine, "Operations for learning with graphical models," *Journal of Artificial Intelligence Research*, vol. 2, pp. 159–225, December 1994.

[18] J. K. Kruschke, *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*, 2nd ed. Bloomington, IN: Elsevier, 2015.

[19] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. Boca Raton, FL: CRC Press, 2013.

[20] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2008, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org