Measuring User Engagement

Alex Tairbekov

Many social media sites have this pattern:

- User A makes a post on the site.
  - User B writes a comment on User A’s post.
  - User A writes a reply to User B’s comment.
  - User B replies back to User A.
  - User C can also comment on the post itself or reply to any of the comments.
- At some point, users stop interacting with the main post.
- What we want to do is find out how likely a user will be to keep a conversation going by replying back to someone’s reply.
- The frequency of a user’s reply is an important factor, and can be described by a stochastic process.

Stochastic Process

- A **stochastic process** is a collection of random variables defined on a probability space, indexed by some **index set** T.
- Some important types of stochastic processes are:
  - **Poisson process**: Each value is independently timed.
    - Example: time between someone's reply to your post and your reply back.
    - Generally works for short intervals of time, but can be disrupted at any moment.
  - **Markov process**: The next value of the variable depends on the previous value only.
    - Example: same as above, but taking into account reply content and time of post.
    - This can account for changes such as the time people are available to post.
- In general, each user will have a different stochastic process for replying to each other user: assume that a directed edge has its own process.

Example: Analyzing user posts on reddit.com

- We can analyze how responsive a given reddit user is by looking at how frequently they enter a conversation with another user.
- In this particular situation votes are ignored and we are solely concerned with the frequency of user interaction - how chatty a user is.
- Sample locations to analyze posts from:
  - Media: /r/books, /r/movies, /r/art
  - Scholarly sources: /r/math, /r/compsci, /r/history
  - Programming: /r/programming, /r/cpp, /r/Python
  - Question boards: /r/AskReddit, /r/AskScience
  - Reddit search results of specific terms (such as “stochastic process”)

File Edit View Search Terminal Help

Areas of mathematics interesting/useful for a theoretical physicist? https://www.reddit.com/r/math/comments/xtkvp/areas_of_mathematics_interestinguseful_for_a/?ref=search_posts
Differential Equations with uncertainty? https://www.reddit.com/r/learnmath/comments/wqqic/differential_equations_with_uncertainty/?ref=search_posts
Want to head into the quantitative finance sector, not sure if I should do a financial mathematics masters, or econometrics. Help! https://www.reddit.com/r/math/comments/w52a4/want_to_head_into_the_quantitative_finance_sector/?ref=search_posts
EURO 2012 - Probability for each team to reach the final round https://www.reddit.com/r/soccer/comments/v56ke/euro_2012_probability_for_each_team_to_reach_the/?ref=search_posts
Accurate Estimation Methods with Small Sample Size https://www.reddit.com/r/statistics/comments/v3oe2/accurate_estimation_methods_with_small_sample_size/?ref=search_posts
good text books about stochastic and stochastic processes? https://www.reddit.com/r/Physics/comments/uw11q/good_text_books_about_stochastic_and_stochastic/?ref=search_posts
My Kaggle rank is leveling out. Advice? https://www.reddit.com/r/MachineLearning/comments/utjdl/my_kaggle_rank_is_leveling_out_advice/?ref=search_posts
Precise example of a martingale? https://www.reddit.com/r/probabilitytheory/comments/uhw0h/precise_example_of_a_martingale/?ref=search_posts
Coupled Stochastic Differential Equations. https://www.reddit.com/r/math/comments/tzf1z/coupled_stochastic_differential_equations/?ref=search_posts
[Stochastics] Conditional Distribution of Poisson Process question https://www.reddit.com/r/cheatatmathhomework/comments/t382v/stochastics_conditional_distribution_of_poisson/?ref=search_posts
Engineering math minor? https://www.reddit.com/r/EngineeringStudents/comments/snxvl/engineering_math_minor/?ref=search_posts
Which is more useful for OR, real analysis or graph theory? https://www.reddit.com/r/sysor/comments/sgpbz/which_is_more_useful_for_or_real_analysis_or/?ref=search_posts
Interesting little probability question https://www.reddit.com/r/math/comments/ryy5n/interesting_little_probability_question/?ref=search_posts
Another Advice Question for Classes https://www.reddit.com/r/math/comments/rodcb/another_advice_question_for_classes/?ref=search_posts
Why is a geometric Brownian Motion with with constant drift and volatility suitable to model the prices of options? https://www.reddit.com/r/math/comments/rcqqz/why_is_a_geometric_brownian_motion_with_with/?ref=search_posts
Question about sample correlation between two stochastic processes. https://www.reddit.com/r/statistics/comments/r6ewu/question_about_sample_correlation_between_two/?ref=search_posts
Backward Stochastic Differential Equations and Feynman-Kac Formula for Multidimensional Levy Processes, with Applications in Finance https://www.reddit.com/r/quant/comments/r5rqj/backward_stochastic_differential_equations_and/?ref=search_posts
Does anyone on r/math have a suggestion for a good stochastic process book for an undergrad? https://www.reddit.com/r/math/comments/qttvr/does_anyone_on_rmath_have_a_suggestion_for_a_good/?ref=search_posts
Suggestions for classes. https://www.reddit.com/r/math/comments/qkrje/suggestions_for_classes/?ref=search_posts
List of Applications https://www.reddit.com/r/math/comments/qc2me/list_of_applications/?ref=search_posts
Good sources for learning signal processing and stochastic process mathematics? https://www.reddit.com/r/learnmath/comments/qagxh/good_sources_for_learning_signal_processing_and/?ref=search_posts
A discussion on free will [self] https://www.reddit.com/r/atheismbot/comments/pgggc/a_discussion_on_free_will_self/?ref=search_posts
A discussion on free will https://www.reddit.com/r/atheism/comments/pgeok/a_discussion_on_free_will/?ref=search_posts
Textbook suggestions? Probability spaces, sigma-algebras, filtrations, measurable random variables, stochastic processes, and martingales https://www.reddit.com/r/math/comments/pg6x5/textbook_suggestions_probability_spaces/?ref=search_posts
Quick question about Gaussian processes https://www.reddit.com/r/statistics/comments/pb91t/quick_question_about_gaussian_processes/?ref=search_posts
[Programming Challenge] Poker Texas Hold'em Simulation and Stat Tracker https://www.reddit.com/r/programingchallenges/comments/ouxbg/programming_challenge_poker_texas_holdem/?ref=search_posts
[Request] library.nu - textbook https://www.reddit.com/r/trackers/comments/odd5f/request_librarynu_textbook/?ref=search_posts
Stochastic Processes https://www.reddit.com/r/math/comments/nbldt/stochastic_processes/?ref=search_posts
How are putative "mental modules" put forward by evolutionary psychologists specified at the neural or genetic level? https://www.reddit.com/r/askscience/comments/m6atb/how_are_putative_mental_modules_put_forward_by/?ref=search_posts
How are problems graded? https://www.reddit.com/r/aiclass/comments/m4g4g/how_are_problems_graded/?ref=search_posts
Won't all Stochastic environments be inherently Partially Observable ? https://www.reddit.com/r/aiclass/comments/leyn7/wont_all_stochastic_environments_be_inherently/?ref=search_posts
A hypothetical intelligent agent: partially observable, discrete, stochastic, non-adversarial? https://www.reddit.com/r/aiclass/comments/ldstv/a_hypothetical_intelligent_agent_partially/?ref=search_posts
AI and Spirituality (Stochastic, Episodic, Continuous, Fully Observable) https://www.reddit.com/r/aiclass/comments/lcuht/ai_and_spirituality_stochastic_episodic/?ref=search_posts
Expected time for probabilistic iterative process to terminate https://www.reddit.com/r/math/comments/kzh1v/expected_time_for_probabilistic_iterative_process/?ref=search_posts
Apologies if this question is too technical, but is there any real life examples of Stochastic Differential Equations? https://www.reddit.com/r/askscience/comments/kx4pj/apologies_if_this_question_is_too_technical_but/?ref=search_posts
Feedback on a scene introducing my villain https://www.reddit.com/r/writing/comments/krw56/feedback_on_a_scene_introducing_my_villain/?ref=search_posts
best/most interesting intro OR course? https://www.reddit.com/r/sysor/comments/kfvjk/bestmost_interesting_intro_or_course/?ref=search_posts
How can I get into more theoretical math? https://www.reddit.com/r/learnmath/comments/kaqre/how_can_i_get_into_more_theoretical_math/?ref=search_posts
Looking for help in stochastic processes course. https://www.reddit.com/r/math/comments/k7gnu/looking_for_help_in_stochastic_processes_course/?ref=search_posts
Double majoring in statistical science or financial math/statistics to help me as Computer Scientist https://www.reddit.com/r/compsci/comments/k12le/double_majoring_in_statistical_science_or/?ref=search_posts

Intended results:

- Each user has their own stochastic process for each edge that they interact with.
- These edges can be summarized into the user's personal process.
- Using this process, we can try to predict how likely a user is to be responsive.
- See https://github.com/alextairbekov for detailed results on May 18, 2017.