NAME: Nicholas Hanoian
DATE: May 4, 2020
Homework: Final Project – Burstiness and Memory in GitHub Activity Data
STAT/CS 387

1 Introduction

When one task is too difficult or time consuming to be done by one person, people form a team and work together to achieve a common goal [1]. This is not just important for everyday life, but also for businesses and their employees. Teams have been recognized as critical to business success for some time [2]. It is important that people not only form teams, but also that these teams are effective. Much work has been done to develop methods to help team members work more effectively on their own as well as with the group [3]–[5]. Some methods such as agile software development and the Scrum framework place a particular focus on working rapidly to complete small goals. Many focus on qualitative studies of teamwork and success, but do not gain the insights that are inside the vast data available [6].

With modern developments in computational power and the increasing availability of large datasets, researchers are now able to apply quantitative methods to the study of teams, not just relying on small case studies [7]. People are even using technology to put together large teams to solve problems in mathematics [8] and computer science [9]. We are particularly interested in the temporal patterns which emerge from how people do work. Anecdotally, we would imagine that people do work in a fairly "bursty" manner. Someone might work for a few days on a project, not touch that project for several weeks, and then work on it for a few more days on end. This idea of burstiness is, in fact, quantified researchers which allows us to analyze whether people actually do work in a bursty manner, as well as the bearing it may have on their success [10]. Alongside burstiness is another measure as well, called the memory coefficient, and this value measures the degree to which the time since the last event influences the time until the next one. These metrics have been used by many researchers to study such things as edit wars on Wikipedia and gene activity in E. coli [11]–[13]. Bursty signals have been observed in many types of humna behavior including emails [14], library loans [15], printing [16], and telephone calls [17].

GitHub is now the leading online platform for open source software development. For public projects, all activity related to the project is publicly available, including when contributions are made, when team members are added, and when general users interact with the repository. In this study, we build off of the work of [18] by utilizing GitHub activity data to analyze whether people work in a bursty manner. We present the results of analyzing more than 150,000 teams, looking not only at whether teams exhibit burstiness, but also whether burstiness is correlated with success. This gives insights into ways to structure teams so that success can be maximized.

2 Methods

2.1 Dataset and team selection

Our dataset is a reproduction of the one used by Klug and Bagrow in [18]. It is data about activity on GitHub ranging from January 1, 2013 to April 1, 2014. Utilizing data cached by GitHub Archive (https://www.gharchive.org/), we aggregated information about different repositories, or teams, and users. This aggregated dataset was used to calculate all of the variables from the original paper: success S, team size, effective team size, total work, experience, diversity, number of leads, and age. We use the same metric of success as Klug and Bagrow which is the maximum number of stargazers the repository obtained during the data window. Using these variables, we reproduced all of the figures from the original paper. The reproductions and copies of the original figures can be found in the Appendix. Some of the figures have more than minor differences because of ambiguities in the original paper by Klug and Bagrow.

The aggregated data was then filtered according to what Klug and Bagrow specified in [18]. Significant teams are defined to be those who had at least one stargazer during the data window and which had at least two updates per month on average (i.e. at least 30 total pushes). After applying these filters, we arrived at 150,039 teams, which is around 1 percent fewer teams than what Klug and Bagrow found originally. This is likely due to some of the slight ambiguities found from following the work of Klug and Bagrow.

2.2 Interevent distributions

Interevent time distributions of complex systems have been studied by researchers in many disciplines including internet activity [19] and seismic activity [20], [21]. We use $P(\tau)$ to denote a distribution of interevent times. An interevnet time distribution with random interevent times is a Poisson process and will follow an exponential distribution. However, if interevent times are not random, the distribution will not be exponential. Goh recommends using either the Weibull distribution (sometimes called the scaled exponential distribution) with probability density function

$$P_{\mathrm{W}}(\tau) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} e^{-(x/b)^a},$$

or the log-normal distribution [10] with probability density function

$$P_{\mathrm{LN}}(\tau) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right),$$

both of which have been used by researchers in the past to model interevent time distributions [22], [23]. Others also consider power law distributions for some datasets [24].

We considered various ways of grouping events in order to derive interevent times. One possible method would be to only consider interevent times between pushes by the same user in the same team. Analyzing the distributions for each member of a team would be asking whether team members in a bursty manner. Another method would be to only require that events come from the same user, ignoring what team that event pertained to. This would be looking at the patterns of human work in general. What we decided to consider was interevent times between pushes by people on the same team. We consider the team to be the atom, not the team member. We choose to investigate the question from this perspective because it keeps things simple. We do not need to worry about small amounts of activity by some members of a team, and we stay centered around the analysis of teams.

Given a set of interevent times, the parameters of these distributions can fit using maximum likelihood estimation [25]. We utilize `fitdistrplus`, an R package, to fit both of these distributions as well as exponential distributions [26]. Figure 1 shows the results of fitting each of these distributions to six different repositories' interevent times. The solid dark gray line shows the empirical distribution, and the color lines represent the three different distributions. The insets in each subfigure also show the time series of each repository's activity. With some of the repositories, both the log-normal and Weibull distributions fit quite well (a, d, d, f), but with others (b, c) the distributions offer very poor fits. The exponential fits are also included to show what a random interevent distribution would look like.

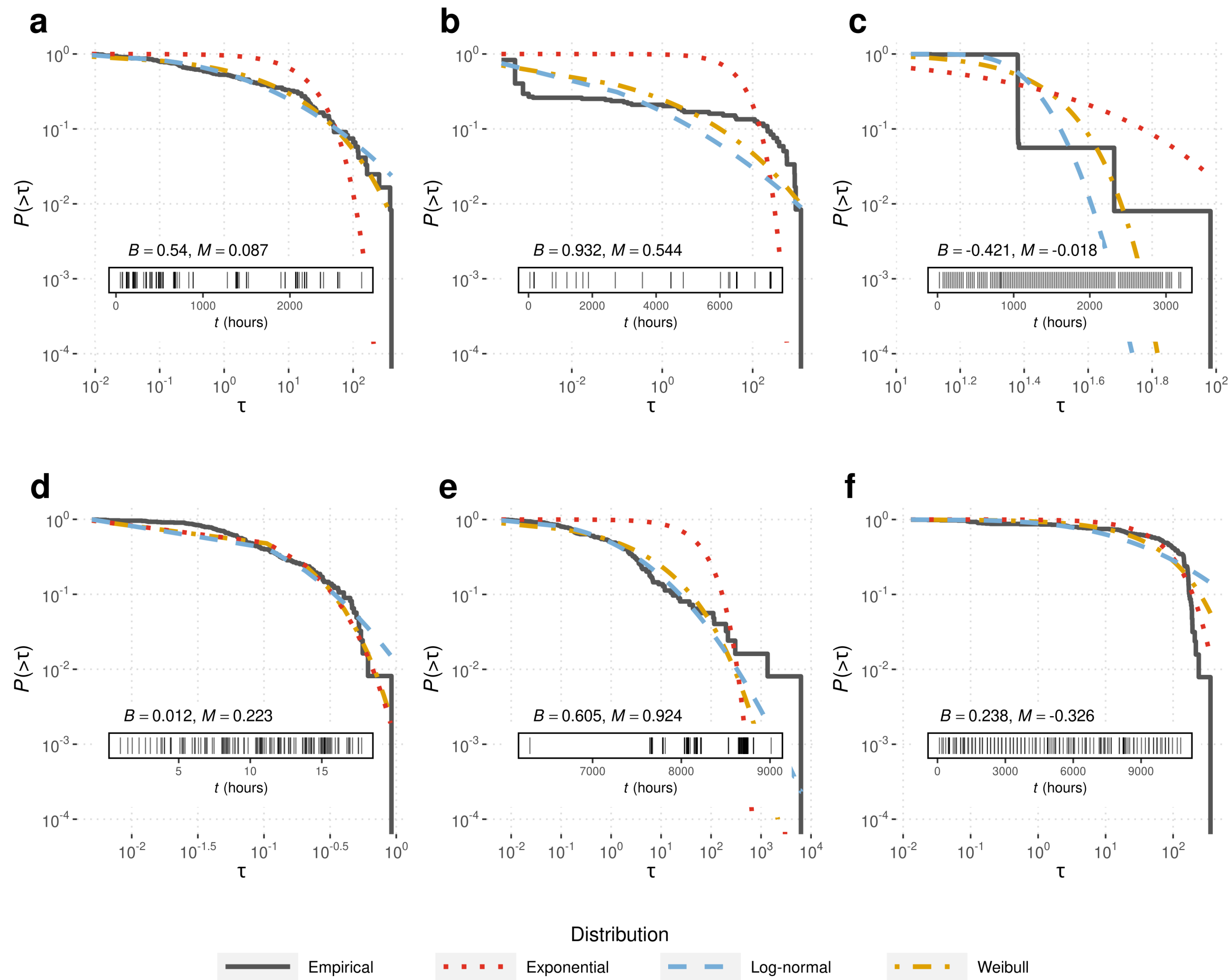

Figure 1: Empirical, exponential, log-normal, and Weibull complementary cumulative distributions for three different repositories. The inset shows the times of commits each repository. Values of B and M are also shown for each repository. The six repositories shown are (a) a repository with B and M around μ_B and μ_M, (b) a repository with very positive B, (c) a repository with very negative B, (d) a repository with approximately 0 B, (e) a repository with a very positive M, and (d) a repository with a very negative M. All repositories shown have approximately the same number of pushes (between 120 and 130), chosen such that their number of pushes is close to the mean number of pushes of all repositories (125.8).

2.3 Burstiness B

The burstiness parameter is defined in terms of the coefficient of variation of the distribution $P(\tau)$

$$B \equiv \frac{\sigma_\tau/m_\tau - 1}{\sigma_\tau/m_\tau + 1} = \frac{\sigma_\tau - m_\tau}{\sigma_\tau + m_\tau},$$

where σ_τ is the standard deviation of $P(\tau)$ and m_τ is the mean of $P(\tau)$. For $P(\tau) = P_{\mathrm{W}}(\tau; a, b)$,

$$m_\tau = b\Gamma\left(1 + \frac{1}{a}\right) \qquad \sigma_\tau = \sqrt{b^2\left(\Gamma\left(1+\frac{2}{a}\right) - \Gamma\left(1+\frac{1}{a}\right)^2\right)},$$

where $\Gamma(\cdot)$ is the standard gamma function. For $P(\tau) = P_{\mathrm{LN}}(\tau; \mu, \sigma^2)$,

$$m_\tau = \exp\left(\mu + \frac{\sigma^2}{2}\right) \qquad \sigma_\tau = \sqrt{\left(\exp\left(\sigma^2\right) - 1\right)\exp\left(2\mu + \sigma^2\right)}.$$

The parameter ranges from -1 to 1, with negative values indicating anti-bursty behavior, 0 indicating random behavior, and positive values indicating bursty behavior. For example, in Figure 1 (a) is moderately bursty, (b) is extremely bursty, and (c) is moderately anti-bursty. As can be seen in the insets of these subfigures, repositories with several tightly-packed activity chunks are very bursty, and those with evenly spread out events are anti-bursty.

2.4 Memory coefficient M

The memory coefficient is essentially the perason correlation coefficient between $\tau_{1,\ldots,n_\tau-1}$ and $\tau_{2,\ldots,n_\tau}$. It is defined as

$$M \equiv \frac{1}{n_\tau - 1}\sum_{i=1}^{n_\tau-1}\frac{(\tau_i - m_1)(\tau_{i+1} - m_2)}{\sigma_1\sigma_2},$$

where $i = 1, 2, \ldots, n_\tau$ are the indices of τ.

M, like B, ranges from -1 to 1, with positive values indicating that short interevent times tend to be followed by a short one, and negative values indicating that short interevent times tend to be followed by long ones. Examples of different values of M can be seen in Figure 1. Subfigures (a) and (c) have M close to 0, (b) and (d) have moderately positive values of M, (e) has a very positive value of M, and (f) has a fairly negative value of M.

3 Results

3.1 Choosing $P(\tau)$

We began by deciding whether to use $P_{\mathrm{W}}(\tau)$ or $P_{\mathrm{LN}}(\tau)$ to model our data. Without considering the data, the Weibull distribution is preferred by default because it has the useful property of reducing to an exponential distribution when $a = 1$ [10]. This makes intuitive sense for the model because when $a = 1$, the value of B is 0. Figure 2 (a) shows the distribution of all of the interevent times pooled together. Visually, the Weibull distribution fits the best, especially in the tail which is where it is hardest to find a good fit. The exponential is far from a good fit, indicating that people's activity on GitHub is inherently bursty and not random. The log-normal distribution is very similar to the Weibull up to around $\tau = 100$, but after that it does not fit the data well at all, expecting many more extreme values in the tail than what actually show up. However, if we look at a goodness of fit measure such as the AIC for each of the distributions, we see contradictory evidence: the AIC for the log-normal model is 97,296,765 while the AIC for the Weibull model is significantly higher at 98,970,588. Using these AIC values, the likelihood of the model given the data is

$$\mathscr{L}(g_i|x) \propto \exp\left(-\frac{1}{2}(\text{AIC}_i - \text{AIC}_{min})\right),$$

as shown by [27]. Calculating this for our two models, we find a likelihood of 1 for the log-normal model and a likelihood of 0 for the Weibull model.

Continuing on our task to determine the best distribution, we fit all three distributions to every team's activity data. Using each of these we can calculate the burstiness parameter B_{W} for the Weibull fits and B_{LN} for the log-normal fits. Now looking at Figures 2 (b) and (c) there are the distributions of these two statistics. The distribution of B_{W} is much wider, yielding a much more diverse range of burstiness values. The distribution of B_{LN} in contrast has an extreme peak at $B_{\text{LN}} = 1$ and a very small tail covering other values. The Weibull distribution is preferred in this regard.

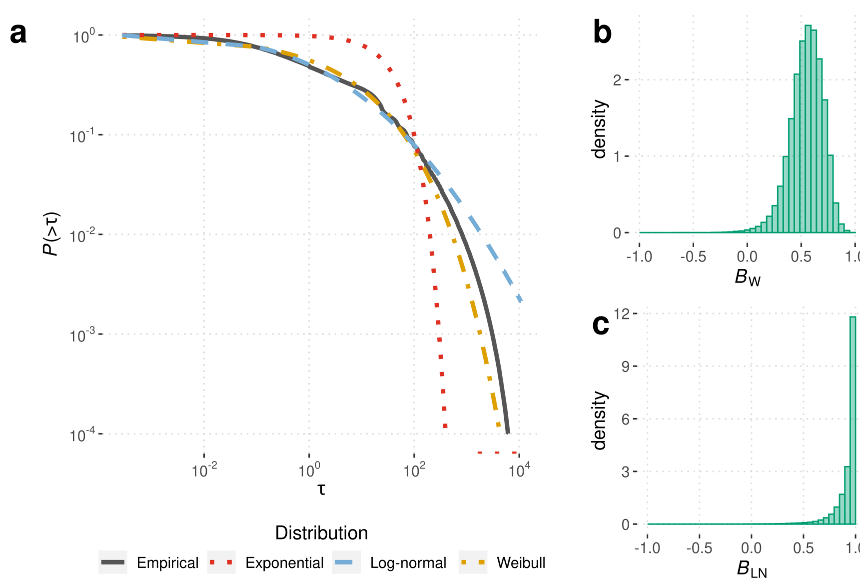

Figure 2: (a) Empirical, exponential, log-normal, and Weibull complementary cumulative distributions for interevent times from all teams pooled together. The Weibull distribution appears to provide the best fit, especially in the tail. (b) Distribution of B_{W}, burstinesses parameters for all repositories calculated using a Weibull distribution to model $P(\tau)$. (c) Distribution of B_{LN}, burstinesses parameters for all repositories calculated using a log-normal distribution to model $P(\tau)$. The distribution of B_{W} is much more spread out, offering a diverse range of values to represent differing levels of burstiness. The distribution of B_{LN} on the other hand is very close to 1 for most repositories.

Just as with the pooled distribution of interevent times, we can use the AIC values for the two fits for every team to see which models the data better. We find that for 104,406 teams (69.6 percent of teams) the log-normal model has a lower AIC value, and for the other 45,630 teams (30.4 percent of teams), the Weibull model has a better AIC value.

Hence, in both visual analyses, the Weibull model fits better, and with both goodness of fit measures, the log-normal model fits better. Due to the fact that the Weibull model fits the tail of the pooled data so much better, and that the distribution of B_{W} is much more descriptive than that of B_{LN}, we choose to use the Weibull model for the remainder of our analyses. In addition, the fact that [10] uses the Weibull model for their analyses pushes us further towards the Weibull distribution.

3.2 Joint distribution of B and M

Now that $P(\tau)$ has been selected, there is now a single B statistic for each team as well as a single M statistic. Figure 3 (a) shows a heatmap of the joint distribution of B and M for all repositories. We see that the distribution is unimodal, and the mode is at $(M = -0.028, B = 0.585)$ (Table 1). If we let this mode summarize all of the teams, then it puts GitHub team activity extremely close to that of human printing activity [10]. It is also close to email patterns and less close to the other human activities that Goh investigates of call center records, phone records, and library loans. In addition, this matches with results of player actions in online games **rmyglod2015interevent**.

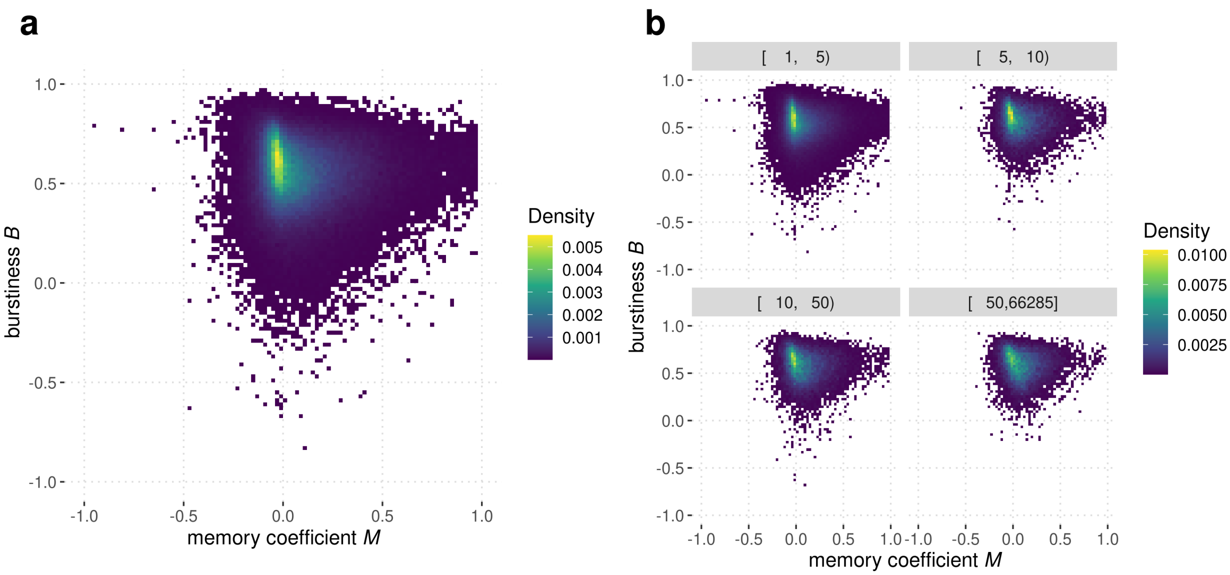

Figure 3: (a) Overall joint distribution of B and M for all repositories. The center of the joint distribution is located at $(-0.028, 0.585)$, shows that GitHub activity data has similar patterns of burstiness and memory as printing data [10]. (b) Joint distributions of B and M for groups of repositories partitioned by S. The range of S values is shown above each subplot. There is little noticable change between subplots, indicating that there is not a strong association between B, M, and S.

3.3 Relationship of B and M with S

We also want to see how B and M relate to S. Figure 3 (b) again shows the joint distribution of B and M, but this time divided into different groups formed by S. The cuts were chosen as a balance between meaningful values of S and keeping enough teams in each group. At a simple visual inspection, the distributions look incredibly similar. However, we can learn more by seeing if the mode of the distribution changes as the values of S under consideration change. The results of using kernel density estimation to find the mode of each distribution are shown in Table 1. We see that there is a clear positive trend in the mode of burstiness, steadily increasing from 0.547 to 0.631 as S increases. This indicates that more bursty teams are associated with greater success. There is also a positive trend in the memory coefficient, increasing from -0.028 to -0.005 as S increases. However, these changes in M are very slight and stay very close to 0 for all categories of S.

S group	Percentage of teams	B	M
All	100	0.585	-0.028
$[1, 5)$	37.11	0.547	-0.028
$[5, 10)$	31.05	0.573	-0.023
$[10, 50)$	8.88	0.610	-0.023
$[50, 66285]$	13.02	0.631	-0.005

Table 1: Most likely values of (M, B) for different levels of S, corresponding to the subplots in Figure 3. As S increases, so does B, indicating that more bursty teams are more successful. As S increases, so does M, becoming slightly less negative and approaching 0. Computed by using kernel density estimation to approximate the distribution of B and M, and then taking maximum value of that estimate.

Now we will consider the relationship between B and M with S, one at a time. Figure 4 shows a scatter plot between each of B and M with S. It also includes marginal density plots to see the univariate distributions of each, as well as a smoothing spline to approximate the relationship between the two variables. The inset in each subfigure shows the same plot, but on a logarithmic scale. Outliers have been excluded from the main plots, but remain in the insets. Beginning with Figure 4 (a), we see that as burstiness increases, there is a slight increase in success. The trend is more pronounced in the inset plot with the logarithmic scale. Figure 4 (b) shows the same plot but with M instead of B. The smoothing spline is almost entirely flat in the linear scale plot, and still flat for most of the actual data in the inset.

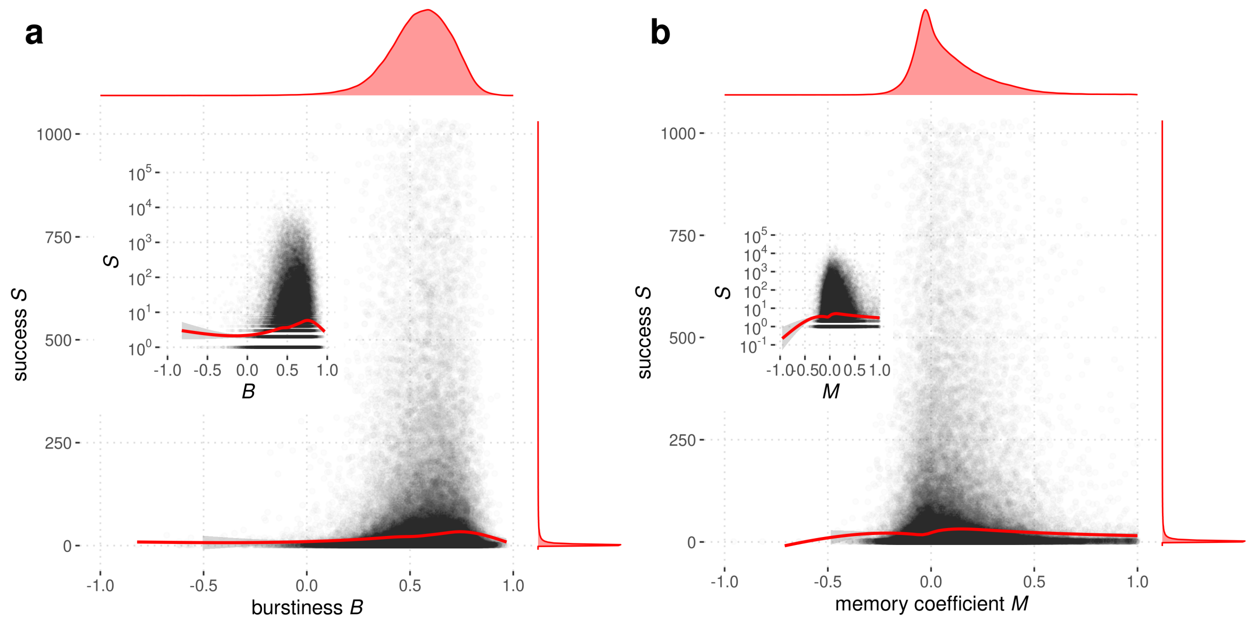

Figure 4: Examining the associations between (a) S and B and (b) S and M. Main plots of each subfigure show a scatterplot of the two variables along with a smoothed spline. Outliers (above the 99th percentile in S) were filtered out of the main plot. Margins show the one-way distribution of each variable. Insets show the same scatterplot and smoothed spline on a logarithmic scale, this time including outliers. There is a slight postive association between S and B, and it is more evident on the logarithmic scale. Based on the splines in (a), higher values of B are associated with higher values of S, to a point. Slightly positive values of M are associated with greater values of S, but the association is minimal.

Finally, we consider the relationship between B and M with S in the presence of the other variables considered by Klug and Bagrow in [18]. An OLS regression model on S was built using all standardized versions of the variables. The estimates for coefficients and associated p-values are shown in Table 2. All variables except for experience are extremely significant. We also fit a model without B or M included, and we find that all of the remaining coefficients are approximately equal in both models. Conducting an ANOVA test to see if the two models are significantly different from each other, we find an F-statistic of 217.52 with a corresponding p-value of less than 2.2×10^{-16}. Hence, B and M do appear to add new information to the model. The coefficients for B and M are both positive, indicating that the more bursty a team is, the more it is expected to succeed, and the more positive the memory coefficient, the more the team is expected to succeed.

variable x	coefficient β_x	p-value
(Intercept)	0.0000 ± 0.0049	1
team size	0.1030 ± 0.0150	0
effective size	-0.1031 ± 0.0124	0
total work	0.0386 ± 0.0050	0
experience	-0.0036 ± 0.0049	0.1485
diversity	0.0370 ± 0.0064	0
no. leads	0.1158 ± 0.0072	0
age	0.1977 ± 0.0050	0
burstiness, B	0.0483 ± 0.0050	0
memory coefficient, M	0.0238 ± 0.0049	0

Table 2: OLS regression model on team success S. Outliers (above the 99th percentile in S) were filtered out to avoid skewing the model. The coefficients for both B and M are highly significant, indicating that they offer valuable information in predicting S even with all other variables in the model. An ANOVA analysis comparing the model with B and M to a model without either found that the models were significantly different ($F = 217.52$, $p < 2.2 \times 10^{-16}$). Note 1: variables are standardized for comparison, so increasing variable x by one standard deviation σ_x implies a $\beta_x \sigma_S$ increase in S. Note 2: p-value of 0 indicates any value less than 2×10^{-16}.

4 Discussion

Teams on GitHub certainly exhibit bursty behavior. The degree of teams' burstiness falls in line with other human activities, specifically printing and emailing. It is important to note that we did not show that people on teams work in a bursty manner, but instead that teams as a whole work in a bursty manner. Because the overall burstiness of the teams falls in line with other human activities, it means that we have shown that groups of people act in similar patterns to individuals. This pattern of burstiness could be explained by *social loafing*, as teams could fall in and out of complacency when many people are working on the same team [28].

When beginning this project, we were hoping to see a much stronger association between particularly burstiness and success than we did. Instead, we found the opposite: teams are able to, for the most part, find success regardless of how bursty or memory-dependent their work patterns are. A more in depth analysis in the future could investigate interevent times grouped by both team and team member. This has its own set of challenges, as more bias would be introduced due to the notable increase in filtering that smaller subdivisions requires. Others have found a strong correlation between burstiness and success in teams, particularly in the context of online crowdsourcing through teams [29].

We are still curious about significant difference in distribution of burstiness parameters depending on which probability distribution is used to fit the interevent time data. Even when a log-norm model is fit to data drawn from an exponential distribution (which should give $B \approx 0$), the value is instead around 0.349. Fitting a Weibull model to the same data performs as expected and yields a value of B very close to 0. Perhaps this is due to an error in our methods, but we have been thorough. Another measure with more desirable properties has been proposed by Kim and Jo, boasting invariance due to finite-size effects [30]. However, it was published in 2016 and does not appear to be gaining much traction

In conclusion, we modeled the interevent time distributions of teams doing work on GitHub, fitting Weibull distributions to each team. Using each fitted distribution's parameters we calculated the burstiness, and we used the interevent data itself to calculate the memory coefficients. We examined the joint distribution of these measures, finding patterns consistent with human behavior data. Finally, we found that there is a very weak, but statistically significant, association between success and burstiness.

References

[1] R. Axelrod and W. D. Hamilton, "The evolution of cooperation", *science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[2] G. M. Parker, *Team players and teamwork*. Citeseer, 1990.

[3] K. Schwaber, "Scrum development process", in *Business object design and implementation*, Springer, 1997, pp. 117–134.

[4] R. C. Martin, *Agile software development: principles, patterns, and practices*. Prentice Hall, 2002.

[5] D. Stokols, K. L. Hall, B. K. Taylor, and R. P. Moser, "The science of team science: Overview of the field and introduction to the supplement", *American journal of preventive medicine*, vol. 35, no. 2, S77–S89, 2008.

[6] M. Hoegl and H. G. Gemuenden, "Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence", *Organization science*, vol. 12, no. 4, pp. 435–449, 2001.

[7] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, *et al.*, "Life in the network: The coming age of computational social science", *Science (New York, NY)*, vol. 323, no. 5915, pp. 721–723, 2009.

[8] T. Gowers and M. Nielsen, "Massively collaborative mathematics", *Nature*, vol. 461, no. 7266, pp. 879–881, 2009.

[9] K. Crowston and J. Howison, "The social structure of free and open source software development", *First Monday*, vol. 10, no. 2, 2005.

[10] K.-I. Goh and A.-L. Barabási, "Burstiness and memory in complex systems", *EPL (Europhysics Letters)*, vol. 81, no. 4, p. 48 002, 2008.

[11] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész, "Universal features of correlated bursty behaviour", *Scientific reports*, vol. 2, no. 1, pp. 1–7, 2012.

[12] T. Yasseri, R. Sumi, A. Rung, A. Kornai, and J. Kertész, "Dynamics of conflicts in wikipedia", *PloS one*, vol. 7, no. 6, e38869, 2012.

[13] L.-h. So, A. Ghosh, C. Zong, L. A. Sepúlveda, R. Segev, and I. Golding, "General properties of transcriptional time series in escherichia coli", *Nature genetics*, vol. 43, no. 6, p. 554, 2011.

[14] J.-P. Eckmann, E. Moses, and D. Sergi, "Entropy of dialogues creates coherent structures in e-mail traffic", *Proceedings of the National Academy of Sciences*, vol. 101, no. 40, pp. 14 333–14 337, 2004.

[15] A. Vázquez, J. G. Oliveira, Z. Dezsö, K.-I. Goh, I. Kondor, and A.-L. Barabási, “Modeling bursts and heavy tails in human dynamics”, *Physical Review E*, vol. 73, no. 3, p. 036 127, 2006.

[16] U. Harder and M. Paczuski, “Correlated dynamics in human printing behavior”, *Physica A: Statistical Mechanics and its Applications*, vol. 361, no. 1, pp. 329–336, 2006.

[17] B. Min and K.-I. Goh, “Burstiness: Measures, models, and dynamic consequences”, in *Temporal networks*, Springer, 2013, pp. 41–64.

[18] M. Klug and J. P. Bagrow, “Understanding the group dynamics and success of teams”, *Royal Society open science*, vol. 3, no. 4, p. 160 007, 2016.

[19] P. Wang, X.-Y. Xie, C. H. Yeung, and B.-H. Wang, “Heterogenous scaling in the inter-event time of on-line bookmarking”, *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 12, pp. 2395–2400, 2011.

[20] S. Touati, M. Naylor, and I. G. Main, “Origin and nonuniversality of the earthquake interevent time distribution”, *Physical Review Letters*, vol. 102, no. 16, p. 168 501, 2009.

[21] M. Bottiglieri, L. de Arcangelis, C. Godano, and E. Lippiello, “Multiple-time scaling and universal behavior of the earthquake interevent time distribution”, *Physical review letters*, vol. 104, no. 15, p. 158 501, 2010.

[22] J. Laherrere and D. Sornette, “Stretched exponential distributions in nature and economy: “fat tails” with characteristic scales”, *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 2, no. 4, pp. 525–539, 1998.

[23] M. Mitzenmacher, “A brief history of generative models for power law and lognormal distributions”, *Internet mathematics*, vol. 1, no. 2, pp. 226–251, 2004.

[24] A. Clauset, C. R. Shalizi, and M. E. Newman, “Power-law distributions in empirical data”, *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.

[25] G. Casella and R. L. Berger, *Statistical inference*. Duxbury Pacific Grove, CA, 2002, vol. 2.

[26] M. L. Delignette-Muller, C. Dutang, *et al.*, “Fitdistrplus: An r package for fitting distributions”,

[27] K. P. Burnham, D. R. Anderson, and K. P. Huyvaert, “Aic model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons”, *Behavioral ecology and sociobiology*, vol. 65, no. 1, pp. 23–35, 2011.

[28] S. J. Karau and K. D. Williams, “Social loafing: A meta-analytic review and theoretical integration.”, *Journal of personality and social psychology*, vol. 65, no. 4, p. 681, 1993.

[29] C. Riedl and A. W. Woolley, “Teams vs. crowds: A field test of the relative contribution of incentives, member ability, and emergent collaboration to crowd-based problem solving performance”, *Academy of Management Discoveries*, vol. 3, no. 4, pp. 382–403, 2017.

[30] E.-K. Kim and H.-H. Jo, “Measuring burstiness for finite event sequences”, *Physical Review E*, vol. 94, no. 3, p. 032 311, 2016.

5 Appendix: Reproductions of figures from Klug and Bagrow [18]

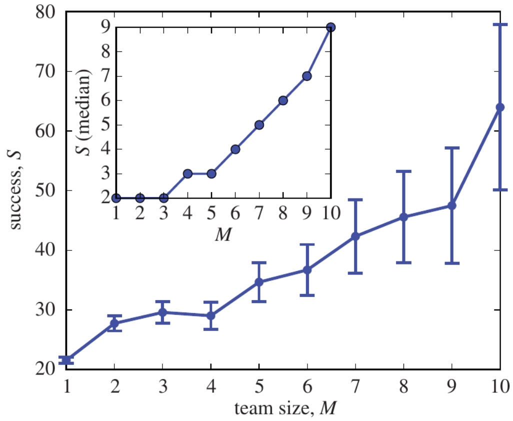

Figure 5: Original version of Figure 1.

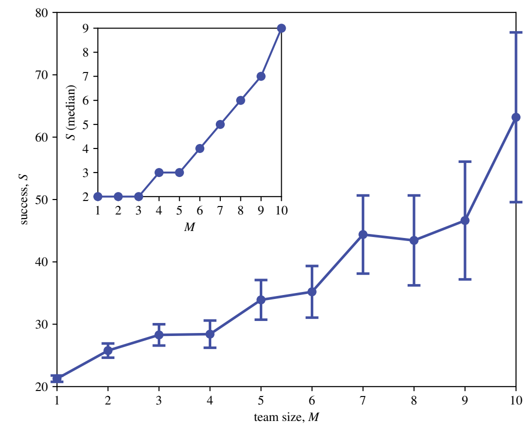

Figure 6: Reproduced version of Figure 1.

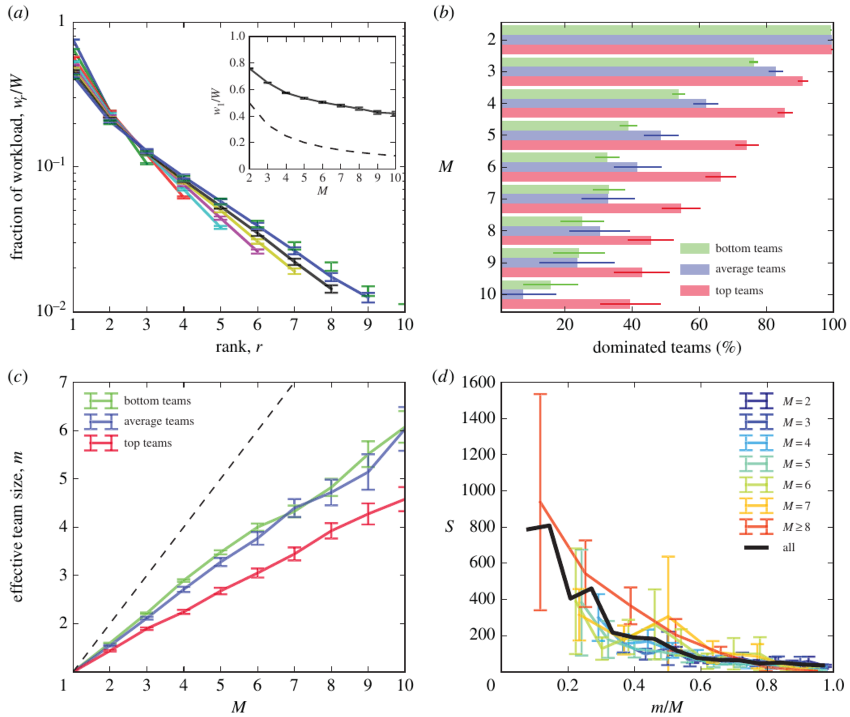

Figure 7: Original version of Figure 2.

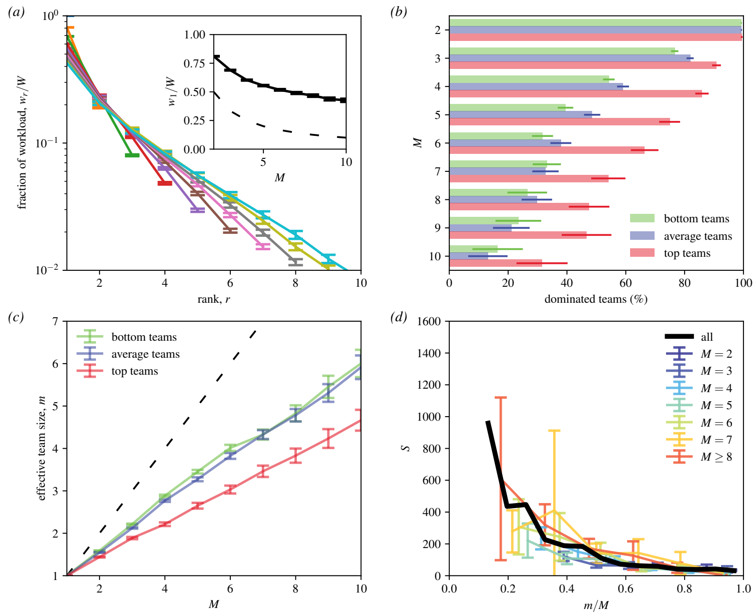

Figure 8: Reproduced version of Figure 2.

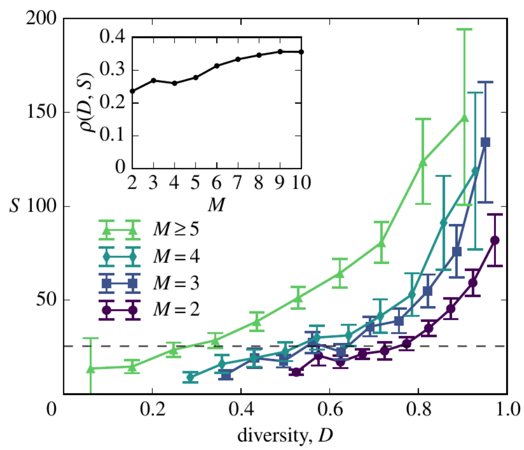

Figure 9: Original version of Figure 3.

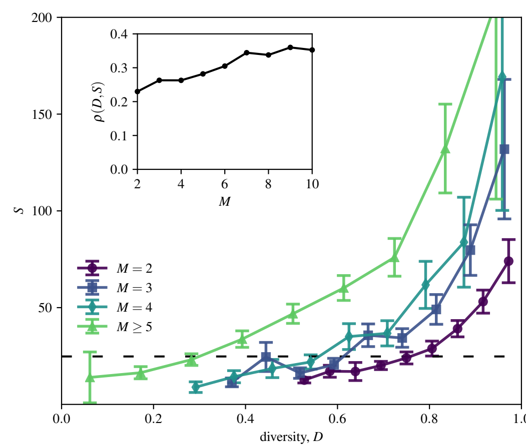

Figure 10: Reproduced version of Figure 3.

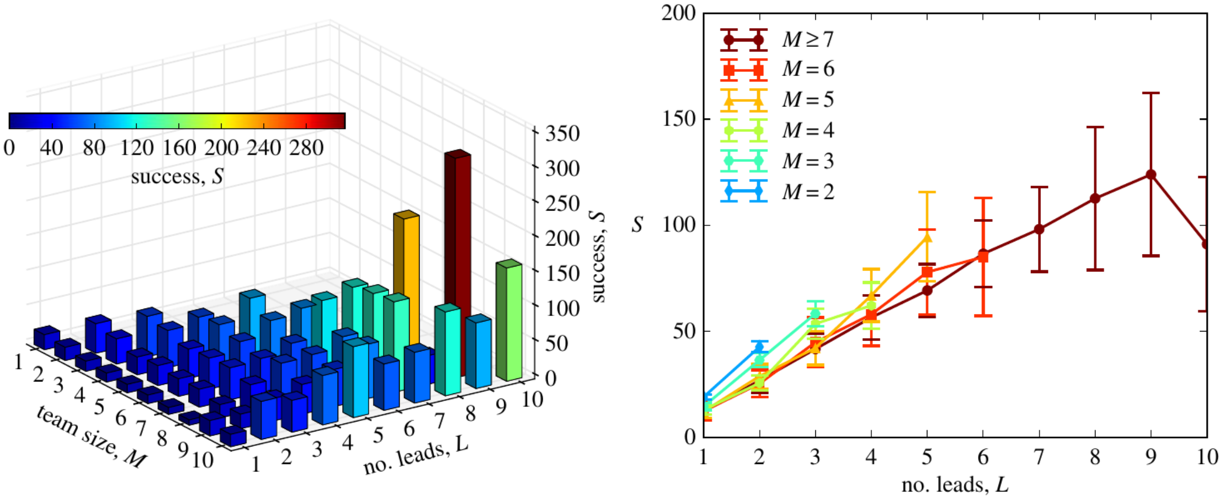

Figure 11: Original version of Figure 4.

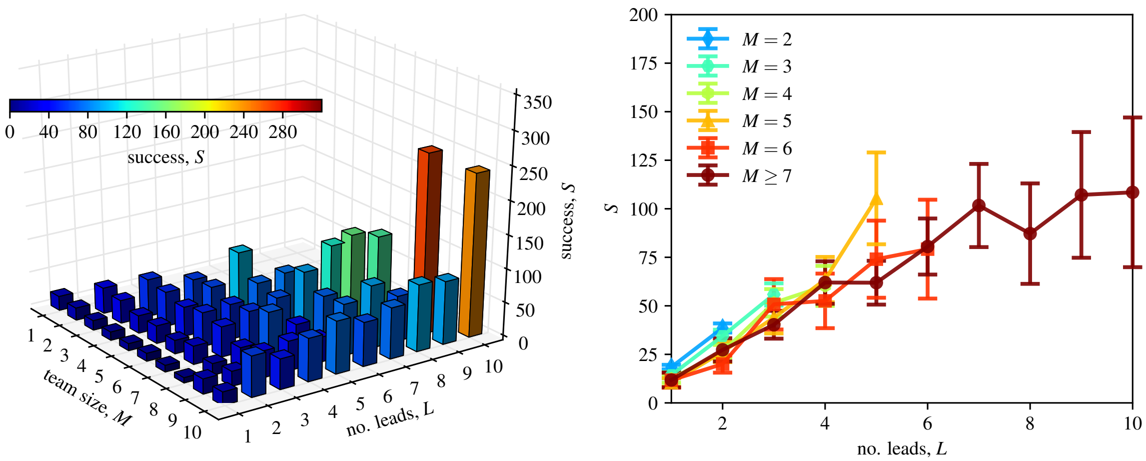

Figure 12: Reproduced version of Figure 4.