Quantitative Research on Network-based Clinical Trials

Prashant Chandrasekar*
Virginia Tech
Blacksburg, Virginia
peecee@vt.edu

ABSTRACT

The Social Interactome (SI) is an ongoing research study of the Addiction Recovery Research Center at VT Carilion Research Institute and the Computer Science and Statistics departments at VT. The high-level aim of the study is to demonstrate through clinical trials involving our closed social-network system whether our proposed interventions aid people in their recovery from addictions. Much of the success of a social-network-based user study that analyzes the behavior of participants, depends on the study design and evaluation. The study design parameters are fine-tuned by instantiating clinical trials, analyzing the data and instantiating the next trial based on the results. So far, much of the research has focused on understanding participant engagement in aggregate forms. Through exploratory research, we aim to analyze the engagement graph to understand the behavior of participant in silo as well as in their friendship groups, to infer network-based characteristics such as homophily, social influence and contagion. The results will be a reinforcement to the efforts of hypothesis testing conducted by the SI team and will have a direct implication in the design of following clinical trials.

KEYWORDS

Social networks, Homophily, Clinical trials, Graph mining

ACM Reference Format:
Prashant Chandrasekar. 2017. Quantitative Research on Network-based Clinical Trials. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, Article 4, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The Social Interactome is an NIH-funded study involving the Virginia Tech Carilion Research Institute (VTCRI) Addiction Recovery Research Center and the departments of Computer Science and Statistics. The project aims to identify scenarios and conditions in which people recovering from substance addiction, placed in a constrained network and provided with self management tools and other aids, are supported in their road to recovery. The study explores ways to select a helpful set of "recovery buddies" who will provide social network support for recovery. One hypothesis on network topology is that selecting sets of recovery buddies with maximal overlap, as opposed to choosing randomly, will be more helpful. Another hypothesis is that homophily-based sets of recovery buddies will be more helpful, i.e., aiding greater engagement and a more sustained recovery. We have run six experiments, with replicates ensuring generality. Each replicate is a clinical trial, where 256 participants, each of whom is recovering from addiction, are recruited and placed in one of two social networks, each containing 128 participants. One network is for the treatment; the other is a control group. Two of the trials included participants recruited from the International Quit and Recovery Registry and ran for 16 weeks , while four of them were for a shorter period and were conducted using Amazon Mechanical Turk. Each of the clinical trials have been driven by hypotheses on the effects of either of network topology, incentive schemes or homophily on participant engagement.

The analyses conducted by the research team has been centered around testing the aforementioned effects, via hypothesis testing. This involves extracting measures, performing aggregation and running p-value based tests. Survival analysis has also been conducted to test the decay in engagement between test and control networks. Furthermore, it is very critical that the network is studied as well. Studying the networks allows us to observe behavioral patterns that would help find meaning behind participants' individual behavior, interactions between pairs and find patterns of use based on certain participant attributes (such as substance in recovery from). This would help us understand common network-related behaviors and help correlate these behaviors against the psychology-related information and other background information that the participants report to us. Through this research, we will extracting different types of interactions, for the same data set, across the latest clinical trial, and try to find some underlying pattern or correlation that could then be used as independent variables for various machine-learning based models or mixed models.

The rest of the paper is structured as follows: In Section 2 we describe the type of data that we collect from each of the clinical trials. We also provide a summary of the various clinical trials and the observations. In Section 3 we cover related work. In Section 4 we outline the aim/goal for this project. In Section 5 we outline what specific observations that we have observed prior to this project. Sections 6 we show our process of data analysis. We end with Section 7 where the implications of our results is discussed.

2 BACKGROUND: DATA

As mentioned in the Introduction section, we have executed 6 clinical trials. Table 1 summarizes some of the important parameters

*

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Table 1: Clinical Trial Parameters

Population	Length of study	Condition to Test
IQRR	16 weeks	Lattice vs. Small-world
IQRR	16 weeks	Lattice vs. Small-world
Amazon Mturk	5 days	Lattice vs. Small-world
Amazon Mturk	12 days	Lattice vs. Small-world
Amazon Mturk	12 days	Lattice vs. Small-world
Amazon Mturk	12 days	Homophily

for the trials. We have used two populations, one recruited from the International Quit and Recovery Registry, a platform where people who are recovering from various forms of addiction voluntarily enroll and engage in self-management tools to help with their substance addiction and the other from Amazon Mechanical Turk. It was due to recruiting challenges that we shifted to using Amazon Mechanical Turk population for the last four studies. Though this might have an effect in terms of correlating engagement and relapse (since the eligibility criteria for Mturkers isn't as stringent), this does not effect this research in studying engagement behavior. For trial one to five, the control group and test group had a fixed topology of either small-world or lattice. For trial six the topology was a small-world network, but the test group had participants assigned friends based on homophily using measures such as primary addiction, age, gender and socio-economic status. For the control group the friends were randomly assigned. The length of the trials were shortened from 16 weeks to 5-12 days because of logistic reasons (we wanted to iterate quickly) and also because we noticed in the 16 week trials, that participant engagement significantly dropped after the first two to three weeks.

The data that we collect from the study fits broadly into three categories:

(1) Web logs: We have instrumented web analytics into the study through the open-source framework, Piwik. Through this framework, we will have a detailed history of all the âĂIJclicksâĂİ by each of the participants, during their visit to the website. This includes when they logged in, every page they visited during their visit, and the length of each of their online sessions.
(2) (For clinical trial 1 and 2 alone) Psychology-related data: The social network interface has components that allow participants to complete surveys that request information about their daily substance use and their social group; additional questions have been created by the psychology âĂIJwingâĂİ of the team. These aim to obtain information deemed important to better understand the circumstances of each participant. Additionally, the participants have the opportunity to attend online group sessions where they engage in a discussion led by a certified moderator. The website also includes testimonials from other people who share their experience of recovery. These testimonials are directly taken from the International Quit and Recovery Registry (IQRR) website. Finally, the participants have the opportunity to view a wide range of âĂIJself-helpâĂİ and/or âĂIJself-managementâĂİ based modules in our Therapeutic Educational System.
(3) Social network related information: During the study, the participants have opportunity to communicate with each other through posts, shares, likes, etc. The interaction can be in textual form and/or through sharing of photos, web links, videos, etc.

Prior to the start of all clinical trials, we have all the eligible participants complete a survey asking their demographic information and substance abuse history.
The web logs and social network information is stored in a MySQL database. The survey information is stored in another server in PostgreSQL.

3 RELATED WORK

One of the big challenges in the analyzing the data from the Social Interactome clinical trials is that a) the network topology is fixed to either small-word or lattice, b) the network is constrained in that each node/participant is connected to no more than 6 nodes/participants with no possibility of adding or deleting connections and c) the participants were monetarily incentivized to engage with one another and therefore not completely organically. With the following network constraints, a lot of the graph-based approaches for learning latent variables such as friendship/trust, attribution for engagement, etc. is challenging. For example, most of the literature pertains to analyzing real-live (open) graphs and observing properties such as path lengths, node neighborhoods, edge-based measures, for the task of link prediction, information diffusion, strength in connection, all of which leads to a wide range of conclusions related to homophily, communities, network topologies, etc. Furthermore, it is important to point out the size of the graph that we are observing and compare that against the graphs have been studied in literature. Therefore, some of the statistical models that have been developed could not be extended or replicated naturally. Given all of the above, the process of literature review was to understand the process of analyses of existing state-of-the-art methodologies. Specifically, the approaches used and the measures aiding the analysis, and understand how we could find similarities in the approach or measures in our dataset. As a result, unlike traditional settings, our research process, along with our aim, is also exploratory in nature. A good amount of literature on social-network mining references the works published in the book Networks, Crowds, and Markets: Reasoning About a Highly Connected World [4]. Chapters describing studies and results about small-world characteristics, homophily and network effects will be appropriately extended to our dataset to extract key measures and approaches to building statistical models. We can use the concept latent spaces as described by Hoff et al., to represent participants in a latent space and observe/infer homophily and observe other attributes or patterns of use for participants that are near to each other in latent space. [5] This model was also used by Sarkar et al. for the task of link prediction by using the distance between two nodes, in the space, as a strong factor in their statistical model [8]. For one of the clinical trails, we constructed a network where each

participant were connected to each other based on homophily, derived from measures such as age, gender, socio-economic status and addiction history. Though not directly applicable, we could learn and use link prediction approaches that have been highlighted in survey paper by Liben-Nowell et al. [7]. Donglei Du presented a summary of techniques for homophily analysis on networks [3]. In the presentation, a threshold-based statistical t-test is proposed as way to observe homophily in networks. The presentation also described research efforts by Kossinet et al. to study affiliation networks to derive triadic closure, and focal closure. [6]. Membership closure (or social influence) was also discussed by describing the approaches used by Crandall et al. and Backstrom et al. [2] [1]. All three approaches define a heuristic and a method to find 1) the probability that two people form a link as a function of the number of "foci" they share and 2) the probability that a person joins a particular focus as a function of the number of their friends already participating in it. For our study, we can transform our data such that the website resources serve as foci and by analyzing the shared interest as well as a time-line of edge formation (indicating interest), we could plot probability distributions on the above metrics.

4 OVERALL AIM

The overall research is focused on properties that relate to participant engagement. Therefore, the goal of this research is to understand the following:

(1) Why did the participants in the homophily-based network engage more than participants who were randomly organized? Was this occurrence random (a coincidence)? If not,
(2) Can we observe and treat homophily as a latent variable and build a graphical model that explains how features of homophily and other factors such as monetary incentives, etc. effect engagement of participants to the website and each other. And finally,
(3) Will this model apply to our observations to the engagement (or lack thereof) that we have observed in our other five clinical trails. If not, what else would we need to consider to make this model more generalizable to our setting.

5 CURRENT OBSERVATIONS

The results of hypothesis testing for the first five clinical trials showed no statically significant difference in engagement and survival analysis between each topology. However, for trial six, the homopholous-based network had statistically more engagement than the non-homopholous-based network. For each of the Mturk based trials, we incentivized participants to engage with one another. For each of these trials, the data showed that incentive schemes worked simply because the high volume in engagement (in comparison to the 16 week trials).

6 ANALYSIS

As described in the Background section, in our most recently conducted clinical trial, E2M4, we have a two sub-networks of participants and have collected their peer-to-peer engagement measures and website usage measure, along with some psychology-based survey information. Our problem is to make sense of the design decisions to help understand the role/weight of each facet of our "intervention/modification", which includes, the factors we have chosen for homophily.

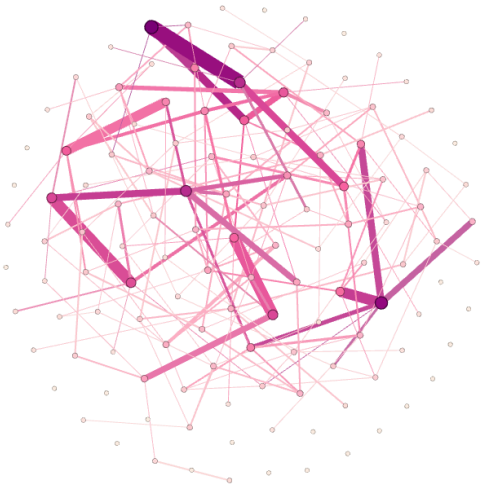

Figure 1: Figure showing peer-to-peer engagement in the homophily sub-network

In this section, we report a step-by-step methodology of our analysis of the homophily-based clinical trial data. As a reminder, the "test network" refers to the homophily-arranged network and the "control network" refers to the randomly-arranged network.

6.1 Explore Peer-to-Peer Engagement Numbers

We start our exploration by looking at the data at a high level and hone-in as required. Figure 1 shows the engagement of participants among themselves in the test network. As you can see, a lot of the edges are no engaged. This was similar to what was happening in the control network as well. The average weighted degree for this network was **6.48** as compared to **4.93** in the control network. The slight difference in degree distribution was unobserved in any of our previous clinical trials.

The peer-to-peer engagement can be broken down into a) Private messages that they could send one another, b) Likes on each other's posts and c) Posts (including wall posts, comments, shares and photo uploads). When comparing the two networks, we notice that only for private messages, do the numbers ***significantly differ***. The average values for each engagement type, between networks, and results from the t-tests can be found in Table 2 and Table 3.

As a disclaimer, this was also previously studied by the larger SI team (and not the work of the team on this project alone).

Next Step: Given that messages alone were significantly different across both networks, our next step was to investigate why this was the case.

Table 2: Average(Mean) engagement numbers by type and by network

Network Type	Mean Posts	Mean Likes	Mean Messages
Test	44.25	177.83	22.167
Control	35.0	130.92	7.50

Table 3: T-test to compare engagement numbers, for each type, between the test and control networks. P-value threshold was 0.05

Engagement Type	P-Value	Result
Messages	**0.007**	**Significant difference**
Likes	0.0986	No significant difference
Posts	0.3247	No significant difference

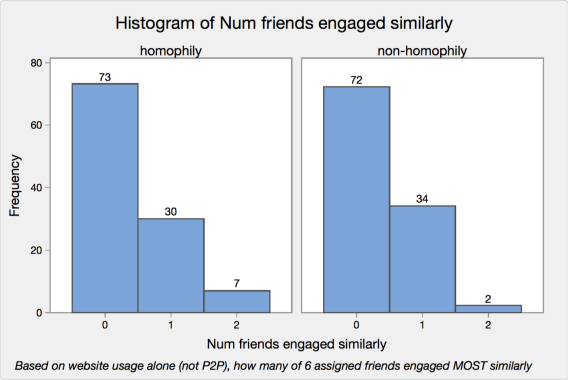

Figure 2: Figure shows that for both networks, based on website usage, how many of their 6 friends engaged similarly

Before that was done, we needed to investigate website resource engagement between networks.

6.2 Explore Website Usage/Engagement Numbers

We wanted to see if homophily was being represented via shared interests on website resources. So we analyzed all the website resources accessed by the participants of both networks, and mapped the top "6" friends that each participant would have based on usage similarity. After that, we would contrast that against the friends given to them based on homophily as defined by personality measures. Figure 2 provided a histogram that shows that both networks, **very few, if not none** of their "assigned" friends shared similar interest based on website resources accessed.

The histogram says that in the homophily network, in 73 cases, none of their assigned friends were in their top six (six because each participant can have only 6 friends), in terms of shared website use.

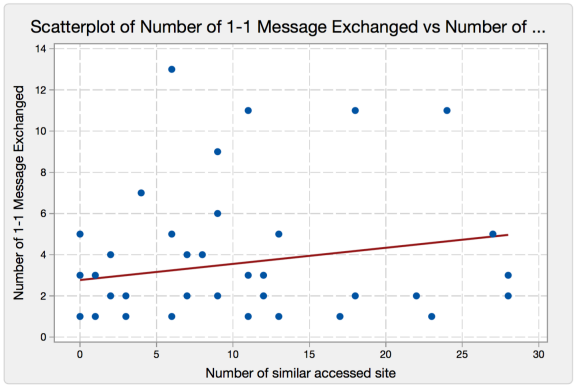

Figure 3: Figure intends to compare engagement between people against their common site usage in the homophily network

This number is comparable to the number for the non-homophily network. This is particular striking, given that we know that by just looking at survival analysis and amount of message engagement, there is statistical difference implying that a homophily-based assignment is driving engagement. But, if we define homophily through pattern of use, it looks like both networks have had random assignment of friends.

To investigate further, we tried to find a pattern or correlate the number of messages sent between a pair of participants against their common website use. From Figure 3, we can see that just because a pair of participants messaged each other a lot, does not necessarily mean that they have a lot in common in terms of interest. The correlation for both the homophily network and non-homophily network was low.

Next Step: Knowing that messages alone is the only differentiating factor, investigate if the features used to define homophily have an effect on the messages exchanged

6.3 Investigating Features of Homophily

As a part of the clinical trial enrollment, each participant has to complete a survey with question that relate to demographics, addiction history and psychology. As subset of these questions are then used as features and a "closeness" metric is determined between every pair of participant. The homophily network is then arranged such that each participant is connected to six of their closest individuals. This particular subset involves 30 questions, that can be answered either by textual input (for questions regarding age, salary, etc.), drop-downs (for example, education, race, etc.), and/or text-box selection (for example, substance in addiction with).

Between each pair of participants, we are trying to understand which particular question or feature in the metric contributes "significantly" to number of messages sent.

6.3.1 Data preparation. We first created a matrix/table with 31 columns and a row for each pair of participants that exchanged at

least 1 message. The 30 columns corresponding to each question in the survey, and the 31st column, the number of private messages exchanged between the pair.

The value for each column was different for each data type provided:

(1) Nominal Values (For ex: age): Difference of the discrete/continuous numbers.
(2) Ordinal Values (For ex: education): Value were first given a logical numbering/order. The difference was the difference between the number.
(3) Categorical (For ex: gender): Since we are representing these values as "differences", if the value of the fields were the same, they would be assigned a value of '0', else '1'.
(4) Selection from text-box (For ex: primary addiction): A symmetric set difference was done on the set of entries for the pair of participants.

We created two such tables/matrices, one for the homophily network and one for the non-homophily network.

6.4 Effect Test via Multiple Linear Regression

A multiple linear regression model was run on the homophily network to see the "fit" and to test the effects. From the 30 "factors", only the following 5 factors had significant effects:

(1) *Factor-7: How many drinks containing alcohol do you have on a typical day when you are drinking?*
(2) *Factor-11: How often during the last year have you been unable to remember what happened the night before because you had been drinking?*
(3) *Factor-20: Please list your preferred or primary substance. e.g drug or alcohol*
(4) *Factor-23: How often do you typically use your second most preferred substance*
(5) *Factor-30: How important is it that a friend with whom you could talk with is the same gender as you*

Following are the specific observations:

(1) Regarding evaluating the fitness, the **R-square adjusted for this model was: 0.3882**. *Note:* In statistics, we have been taught that the interpretation of the R-squared adjusted needs to be contextualized and given relevance, based on the domain. From my understanding and conversation with my colleagues, qualitative responses, when fitted, generally don't produce a high value. Having said that, we cannot use this model for prediction purposes, which is one of the main aims of this research.
(2) Also, regarding the factors, for which the data shows significance, it is good to see primary addiction substance and secondary addiction substance are present, given that they are a central topic for the Social Interactome project.
(3) Given that we have thirty factors, there is always the possibility of collinearity. In our case, it is surprising to see factor-30 being significant, but not the factor pertaining to "difference in gender". Also, factor-11 seems like it *could be* correlated to factor-7, given that they both pertain to alcoholic consumption.

Table 4: Regression Model Comparison

Model	AICc	BIC
Poisson	410.27	284.50
MLR	275.08	285.94
ZI-Poisson	443.10	288.26
Negative Binomial	448,42	293.5
Exponential	447.57	321.8

Next Step: To execute stepwise regression on the same data so that we can avoid collinearity and conduct feature ablation at the same time. We expect the result to also improve the "goodness-of-fit" of the model.

6.5 Stepwise Regression

We conducted a mixed stepwise regression, with our P-value thresholds for entering the model and leaving the model, both being 0.15. Afterwards, based on the factors selected, we ran a multi-linear model.

The following **additional** factors were included:

(1) *Factor-2: Age*
(2) *Factor-5: Yearly Income*
(3) *Factor-12: How often during the last year have you needed an alcoholic drink first thing in the morning to get yourself going after a night of heavy.drinking?*
(4) *Factor-25: WillingnessScore*
(5) *Factor-29: How important is it that a friend with whom you could talk with is the same education as you*

None of the previous factors were excluded.

Using the total **10** factors, the model was run and produced a R-square adjusted value of 0.4661.

Also, and more importantly, only factors 2,7, 11, 12, 23 and 25 were **significant** in this model. (Please refer to the specific definitions/meanings of these factors from above.

To improve our fitness, we also modeled the data using Poisson, Zero-Inflated Poisson, Negative Binomial and Exponential regression models. Table 5, provides the AICc and BIC scores for each of the models.

Next Step: Given that we have identified a subset of factors that have been significant in various models, we want to see how the values of these factors compare with the values in the non-homophily network.

6.5.1 Significant Factor Comparison. From our engagement number, we know that private messages are significantly more in the test network because of homophily. In the previous section, we were able find out features/factors in the homophily calculation that significantly contributed to this. Here we conduct t-tests to compare the values of these factors, one network against another. Please remember, that the factors themselves, represent "a difference" in value for each pair of participants that have messaged each other more than once. Table 5 provides the results.

From the results, the data suggests no statistically significant difference between the factors in both the homophily and non-homophily network.

Table 5: Factor comparison between homophily and non-homophily network

Factor	P-value	Result
Factor-7	0.2350	No significant difference
Factor-11	0.6759	No significant difference
Factor-12	0.9794	No significant difference
Factor-23	0.9232	No significant difference

7 DISCUSSION AND FUTURE WORK

From our analyses above, we observe that though the factors/features of homophily that were significant in their contribution to private messages exchanged, these values **did not** differ from the values in the participant-pair in the non-homophily network. Though this result is not ideal, we can move ahead by looking at sub-cohorts of participant pairs, instead of looking at the data as a whole. As generally seen in a social network, population engages at different levels. We can create various participant-pair cohorts based on different engagement criteria and study them separately. (starting with "high usage participants" for both networks.)

With regards to building a predictive model, though not discussed in the paper, we have run some preliminary models on the effect that monetary incentives has on engagement, more specifically, the independent variable being the amount of money earned on previous day and the dependent variable being the number of actions made on the website. We have also tested the factor of "friendship behavior" on previous day for the model. The work is at an early stage, as the model fitness could be improved on.

We have only been able to address 1 out of the 3 project aims, mentioned in Section 4. This is partly because the other 2 aims are based on the hypothesis that we **will** have the answers/solution for Aim 1. Through our research process, we have come a step closer to understanding why homophily-based network outperformed the random-based network, in terms of engagement (more specifically private messages).

The short-term future work will involve running further models on possibly different cohorts of participants to see if we can highlight features in homophily that manifest in the network. Our long-term work will involve taking these features along with other interventions such as network topology and incentivization schemes to confidently predict when a participant or a participant-pair will engage.

8 CONCLUSION

Through this research, we aim to analyze the reasons for higher engagement in the homophily-based network arrangement as compared to the random-based network arrangement. This would help our broader aim of understanding the role that participant relationships and network play when it comes to providing an environment conducive for participants to engage and support one another in their road to recovery from substance addiction. By identifying factors in homophily that could have an effect, we are able to advise the Social Interactome team during the design of the next clinical trial. With each trial, we hope to discover more, that will have a cascading effect on the design, execution and analysis of further studies.

ACKNOWLEDGMENTS

This research is based on the work towards the NIH-funded project, The Social Interactome of Recovery: Social Media as Therapy Development (NIH Grant 1R01DA039456-01).

REFERENCES

[1] Lars Backstrom, Dan Huttenlocher, Jon Kleinberg, and Xiangyang Lan. 2006. Group formation in large social networks: membership, growth, and evolution. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 44–54.

[2] David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, and Siddharth Suri. 2008. Feedback effects between similarity and social influence in online communities. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 160–168.

[3] Donglei Du. [n. d.]. Social Network Analysis: homophily. ([n. d.]). Retrieved October 15, 2017 from http://www2.unb.ca/~ddu/6634/Lecture_notes/Lecture_6_homophily_handout.pdf

[4] David Easley and Jon Kleinberg. 2010. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press.

[5] Peter D Hoff, Adrian E Raftery, and Mark S Handcock. 2002. Latent space approaches to social network analysis. *Journal of the american Statistical association* 97, 460 (2002), 1090–1098.

[6] Gueorgi Kossinets and Duncan J Watts. 2006. Empirical analysis of an evolving social network. *science* 311, 5757 (2006), 88–90.

[7] David Liben-Nowell and Jon Kleinberg. 2007. The link-prediction problem for social networks. *journal of the Association for Information Science and Technology* 58, 7 (2007), 1019–1031.

[8] Purnamrita Sarkar, Deepayan Chakrabarti, and Andrew W Moore. 2011. Theoretical Justification of Popular Link Prediction Heuristics.. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, Vol. 22. 2722.