

© 2020 IEEE. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Measuring Peer Mentoring Effectiveness in Computing Courses: A Case study in Data Analytics for Cybersecurity

Abu Zaher Md Faridee, Vandana P. Janeja

Department of Information Systems, University of Maryland, Baltimore County
{faridee1, vjaneja}@umbc.edu

Abstract—Computing courses often suffer from lack of diversity. In this paper we evaluate an intervention method of peer mentoring to help increase interest in data analytics in cybersecurity. We present a text mining approach to analyze student assignments while they undergo a peer mentoring exercise. In our prior work, we have shown that the peer mentoring approach is effective at improving the students' interest in cybersecurity careers and contributes to an overall better knowledge gain throughout the semester. This was also reflected by an improvement in grades with two years of anonymous survey results. Across the years we also observed that peer mentoring is particularly effective in diverse groups. In this paper, we perform text mining of the written assignments for analyzing the group behavior of the control and experiment sections of a class while also documenting the effectiveness of intervention methods such as peer mentoring. We employ a few text mining techniques, namely *Text Frequency Analysis*, *Lexical Diversity*, *Readability Analysis*, *Hyperlink usage* and *Objectivity Analysis* on the text assignments submitted by the students and show that students who receive peer mentoring are able to express more complex ideas with fewer words and thereby receive higher grades by the end of the semester. Based on these results, we also discuss how our methodology would be applicable in increasing reachability and diversity in other specialized computing courses such as Big Data and distributed systems.

Index Terms—cybersecurity, peer mentoring, text mining, readability lexical diversity, sentiment analysis

I. INTRODUCTION

The lack of diversity is a major issue in computing careers. Peer mentoring has been shown to increase interest of various levels of students particularly diverse groups. Peer mentoring is a proven intervention method which allows for matching of more experienced students with less experienced students who are typically in the early stages of their programs. Prior work has shown peer mentoring to be effective [1]–[10] in inculcating (a) better academic performance after the first year and reduction of student attrition over all four years in the student mentees and (b) an increase in critical thinking, leadership skills, and abilities for the student mentor. Our framework is based on the evidence from these prior studies indicating that peer mentoring works well in a structured curriculum and shows increased retention with students in the experimental groups earning higher grades and developing critical thinking skills.

We wanted to establish the use of peer mentoring in the cybersecurity curriculum [11], [12]. Thus, in our peer mentor-

ing approach, we pair the undergraduate students from *IS 300: Management of Information Systems* gateway course with the students from (*IS 471: Data Analytics for Cybersecurity*), an advanced cybersecurity course. We choose IS 300 for the peer mentoring as this is usually the place where students begin to explore career pathways and the course offers an opportunity to increase participation in advanced technology courses as the gender and underrepresented groups are still at a slightly higher percentage than in some of our advanced electives. On an average UMBC has about 46% female and 22% underrepresented groups, while the IS department as a whole has roughly 19% female and 22% underrepresented groups. In some of our upper-level electives such as Data Mining (2017 Fall), we have 23% female and 33% underrepresented minorities. In IS 300 we have seen over 27% females and 27% minorities (2017 Spring). Thus, we hope to encourage some of these students into advanced technology careers through peer mentoring and increase participation in this important area of STEM education and cybersecurity careers. We demonstrated in our prior work [11], [12] that students taking introductory courses when engaged in peer mentoring exercises with advanced data analytics in cybersecurity electives expressed more interest in cyber careers. Moreover, peer mentoring was clearly seen to have an impact on the diverse groups, in particular on women. In this paper we extend our work to quantify the level of gains in the peer mentored group through text analysis of the student assignments.

The key contributions of the paper are as follows:

- We primarily utilize text mining techniques to analyze the text assignments written by IS 300 students after the peer mentoring interventions and compare the results between students who attended regular lectures (in a Control section of IS 300) vs students who received peer mentoring (Experimental section of IS 300) for Spring 2017 session.
- We compare the grades associated with the written assignments attained by the students of both groups during two phases of the semester and investigate what properties of the text contribute to higher grades.
- We briefly discuss application of our methodology for increasing reachability and diversity in other specialized courses such as big data and distributed systems.

The rest of the paper is organized as follows: we begin with a

brief outline of the intervention study design in Section II. In section III we evaluate, compare and analyze the trend from the text analysis. Finally, we discuss the future directions of the project in section IV.

II. METHODOLOGY

The methodology of the peer mentoring approach is based on our prior work [11], [12], discussing results from a two semester (Spring 2016 and Spring 2017) offering of this intervention. Figure 1 summarizes the study methodology across the two sections (experiment with intervention: IS 300E and control without intervention: IS 300C) of IS 300 and the interactions with peers from IS 471.

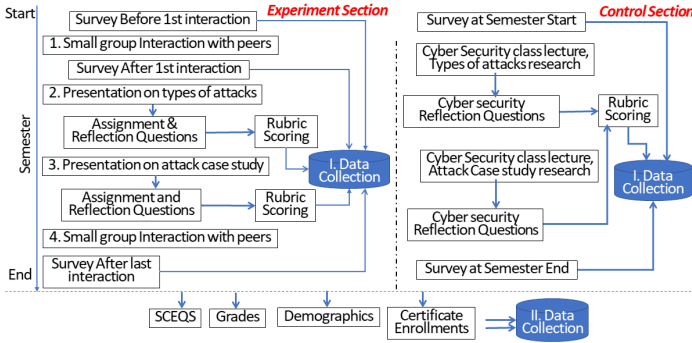


Fig. 1: Study Design: Experiment Vs. Control [11]

A. Peer Mentoring Approach

Once every month, students in IS 300E interacted with the IS 471 students in a classroom setting through small-group conversations and presentations from the advanced class of students. The two courses were scheduled in the same time slot to facilitate coordination. Both IS 300C and IS 300E courses were taught by the same instructor. IS 471 students provided advanced insights, knowledge, and skills to IS 300E students to enhance relevant assignments in cybersecurity. As peer-mentors the IS 471 students also provided information and insights into cybersecurity careers to potentially raise the IS 300 students' awareness of and interest in coursework and career paths in cybersecurity. The interactions between IS 300E and IS 471 included two short interactions (at the beginning and end of the semester) where small groups (with a mix of IS300E and IS471 students) interacted with each other through unstructured discussions, guided by some leading questions regarding cybersecurity careers. Throughout the semester IS 300E students attended presentations on types of attacks and attack case studies done by the IS 471 students. These were quick, short presentations that were highly animated and interactive with lots of Q&A. Following these presentations, students completed short essay type questions on these topics by selecting a type of attack and a case study to research for their assignment. In IS 300C these presentations from peers were replaced by classroom lectures and research done by the students individually. There were two surveys conducted before and after the interventions in the IS 300E and 471 classes. There were also two surveys conducted in the

control section of IS 300 to evaluate the efficacy of classroom instruction about cybersecurity and self-guided assignments.

B. Extending Peer Mentoring Beyond Cybersecurity Curricula

We posit that the peer mentoring model we have proposed and experimented with is adaptable to other types of curricula with presentation elements. As an example let's consider the graduate course, IS 651 Distributed systems, at the Information Systems department at UMBC. This course has a case study element where students present case studies on topics of distributed computing architectures. At the same time we have required core courses such as IS 601: Management Information Systems and IS 603: Decision Technology Systems. For a similarly graded component from IS 300 to IS 471, a component could be modeled for students to explore interests in Distributed systems. This is particularly interesting for undergraduates and graduates alike since UMBC also hosts a new course combining Atmospheric physics with Big Data and High Performance Computing under the NSF Cybertraining efforts for workforce development in these areas¹. Our results across the semesters are consistent that peer mentoring can be supportive for increasing diversity and improving learning outcomes. If the peer mentoring interactions through presentations from upper level courses can be embedded into lower level courses this can be easily adapted to other types of curricula.

III. RESULTS

The gender and ethnicity distribution of the sections across the years 2016 and 2017 are depicted in Table I and II respectively. The total number of students in each group is denoted by 'N' in the table headers. IS 300 students from both control and experiment sections were given two assignments on cybersecurity throughout the semester. The first assignment was on *types of cyber attacks*, where the students were asked to write an essay on a specific type of cyber attack from a predefined list. They were asked to provide a description of the attack, example incidents, commonly affected businesses and preventive measures against the attack. The second assignment was a *cyber attack case study* where the students were asked to write an essay on a recent cyber attack incident. They were asked to provide the background, estimated damage, targeted audience, resolution, and lessons learned from this case study. The difference between the experiment and control section was that the IS 300E students attended presentations made by IS 471 students on multiple *types of cyber attacks* and *case studies* prior to these two assignments. The assignments were graded by the same instructor who taught both the control and the experiment section during that same semester. We discuss the analysis on the grades of the *types of cyber attack* and *cyber attack case study* assignments in Section III-A for both year 2016 and 2017. In the sections that follow, we employ several text analysis and mining techniques on the electronically collected text data of the assignments. In 2016 the data was not collected electronically, so we only show

¹<http://cybertraining.umbc.edu/>

TABLE I: Gender Distribution

Gender	Spring 2017			Spring 2016		
	IS300E (N=39)	IS300C (N=40)	IS471 (N=22)	IS300E (N=38)	IS300C (N=33)	IS471 (N=22)
Male	59%	75%	67%	82%	91%	82%
Female	38%	20%	33%	15%	6%	18%
Unspecified	3%	5%	0%	3%	3%	0%

TABLE II: Ethnicity Distribution (in percentage)

Ethnicity	Spring 2017			Spring 2016		
	IS300E (N=39)	IS300C (N=40)	IS471 (N=22)	IS300E (N=38)	IS300C (N=33)	IS471 (N=22)
Unspecified	8%	5%	6%	3%	24%	13%
White	28%	35%	32%	42%	18%	31%
African-American	15%	17%	16%	16%	15%	15%
Asian	46%	32%	39%	26%	33%	30%
Multiple races	0%	5%	3%	13%	6%	10%
Other	3%	5%	4%	0%	3%	1%

the analysis on the 2017 data where the assignments were electronically submitted.

A. Grade Analysis

On the first *types of cyber-attack* assignment, the IS 300 students from the control section have slightly higher class average (10.97% higher in 2016 and 3.27% higher in 2017) and lower variance than the experiment section, but at the latter part of the semester when they are given the *case study* assignment, they fall behind compared to the students from the experiment section. For the *types of attack* assignment, the control section students may be able to find adequate information over the traditional Internet resources, but the *attack case study* assignment is much more complex in nature, requiring research, analysis and synthesis of conclusions from real world attack case studies. The IS 300E students are possibly able to gain more knowledge here, as they attend the presentations and accompanying Q/A sessions from the IS 471 students. They are able to better understand these complex issues in contrast to the students from the control section who can only use self guided research as the primary resource. This pattern is observable in both 2016 (Figure 2) and 2017 (Figure 3), experiment section's average score increases in the second assignment (7.18% in 2016 and 12.10% in 2017) and

TABLE III: IS 300 Spring 2016 Grade distribution

	Types of Attacks			Case Study		
	Control	Expt.	Change	Control	Expt.	Change
Avg.	9.21	8.20	-10.97%	8.22	8.81	7.18%
Std. dev.	0.54	1.42	N/A	1.72	0.84	N/A

TABLE IV: IS 300 Spring 2017 Grade distribution

	Types of Attacks			Case Study		
	Control	Expt.	Change	Control	Expt.	Change
Avg.	8.65	8.37	-3.24%	8.18	9.17	12.10%
Std. dev.	2.28	2.76	N/A	2.48	1.68	N/A

the variance goes down, the control section exhibits the opposite trend in both years. The results are shown in table III and IV for the Spring 2016 and 2017 sessions respectively.

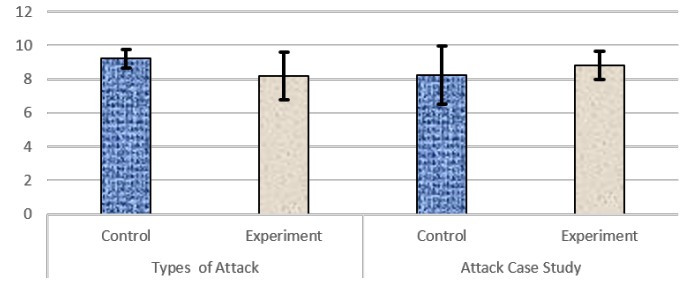


Fig. 2: Comparison of Average Grades in Spring 2016 (black bars show one standard deviation range)

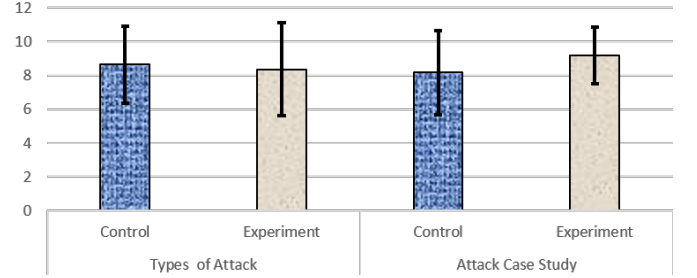


Fig. 3: Comparison of Average Grades in Spring 2017 (black bars show one standard deviation range)

B. Text Frequency Analysis

Before delving into complex text analysis on the assignments, we started off by measuring the word, sentence and syllable usage in them. We took the word, sentence and syllable counts of each of the written assignments which were later compared with their average and standard deviation across the sections. The experiment section constantly shows lower average (with lower variance) word (5.56% and 9.48% lower), sentence (7.22% and 5% lower) and syllable (6.10% and 6.08%) usage (for *Types of Attack* and *Attack Case Study* assignment respectively). The standard deviation of word, sentence and syllable usage across the assignments is also lower compared to the control section. These results are shown in Figure 4, 5, 6. and Table V. These were very interesting findings, as for the second assignment (attack case studies), the students from the experiment section were still able to get 12.10% higher grade while using 9.48% fewer words and 5% fewer sentences. This prompted us to take a deeper look into the composition of the texts with more sophisticated text analysis techniques.

C. Lexical Diversity Analysis

Lexical Diversity is the measure of how many different words are used in a text and it is often used as an equivalent to understanding *Lexical Richness*. The more varied a vocabulary a text possesses, the higher the lexical diversity. For a text to be highly lexically diverse, the speaker or writer has to use many different words, with little repetition of the words

TABLE V: Frequency Analysis of the Assignments in Both Sections

	Types of Attack		Attack Case Study	
	Control	Expt.	Control	Expt.
Avg. Words	396.21	374.17	495.47	448.52
Std. Dev. Words	204.13	176.11	189.63	162.11
Avg. Sentences	27.05	25.10	32.00	30.40
Std. Dev. Sentences	15.62	11.20	13.48	12.30
Avg. Syllables	639.94	600.90	787.45	739.54
Std. Dev. Syllables	345.07	280.43	322.39	277.49

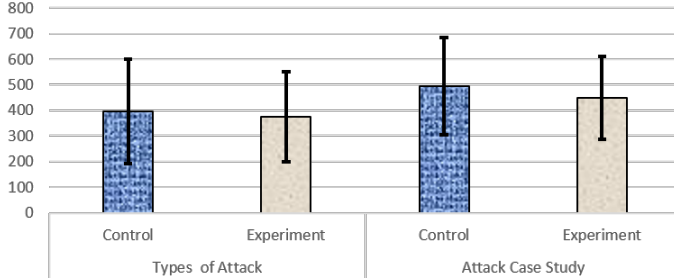


Fig. 4: Mean Word Usage in Spring 2017 (black bars show one standard deviation range)

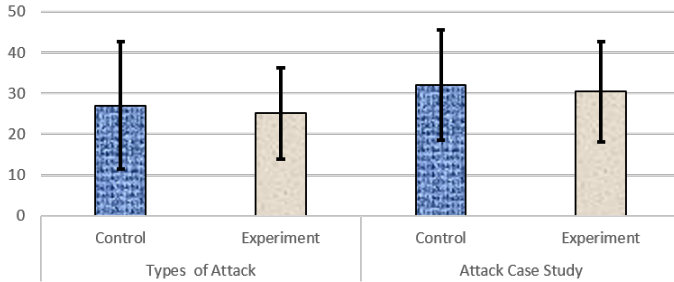


Fig. 5: Mean Sentence Usage in Spring 2017 (black bars show one standard deviation range)

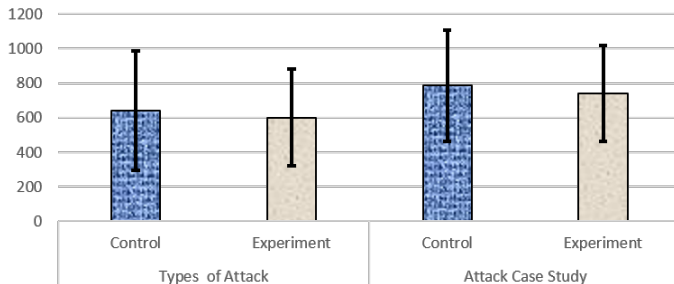


Fig. 6: Mean Syllable Usage in Spring 2017 (black bars show one standard deviation range)

already used [13]. The measurement of vocabulary diversity has often been used in a wide range of educational and linguistic research including child language development, language impairment, foreign and second language learning, the development of literacy, authorship studies, forensic linguistics etc [14]. The traditional lexical diversity measure is the ratio of the different word (types) to the total number of words (tokens) which is often called *Type-Token Ratio* or *TTR*. A common criticism of TTR is that text samples containing large numbers of tokens give lower values for TTR and vice versa;

a longer text, in general, has a lower TTR value than the shorter text. Since all of our assignments are of similar word counts (as we posed a restriction on the maximum length to two pages), we expect this limitation of TTR will not have major effect on our analysis. We also calculate more recent and robust measures, of lexical diversity in our analysis such as *VocD* [14] and *MLTD* [15].

We used *koRpus* [16] package in R to conduct this analysis. The results are shown in Figure 7, 8 and 9 and Table VI. Firstly, the lexical diversity of the *Attack Case Study* assignment is higher compared to the *Types of Attack* assignment for both the control and experiment section. This can be attributed to the nature of the two assignments as the types of attack assignment is more theoretical in nature whereas the case study assignment requires the students to write from real-world incidents. Also, as the semester progresses, the students become more comfortable with cybersecurity nomenclature. In terms of the difference between the control and the experiment section, the experiment section shows more lexical diversity for both assignments with all of the measures. More specifically, For *Types of Attack* and *Attack Case Study* assignment, the *VocD* score for experiment section is 3.96 and 1.65 times higher respectively. The *MLTD* score, on the other hand, is 5.07 and 2.40 times higher for experiment section for *Types of Attack* and *Attack Case Study* assignment respectively. If we take normalized average among all the measures, we can notice that the experiment section has 3.55 and 1.8 times higher lexical diversity than the control section for *Types of Attack* and *Attack Case Study* assignment respectively which is a noticeable difference.

D. Readability Analysis

Readability is the property of some texts being easier to read than others. The creator of the SMOG readability formula G. Harry McLaughlin (1969) [27] defines readability as “the degree to which a given class of people find certain reading matter compelling and comprehensible”. Edgar Dale and Jeanne Chall [28] define *Readability* more comprehensively, as “The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting.”. Flesch-Kincaid Grade Level [29] is one of the most well-known readability formulas which has been extensively used in the field of education. To have a well rounded readability comparison, we calculate several other readability scores (described later in Table VII).

We again used *koRpus* [16] package in R to conduct this analysis. The results are being shown in Figure 10, 11, 12 and Table VII. We observe a similar but much less pronounced phenomenon of what we observed in Section III-C. Between the assignments in the same section (control or experiment), there is a minor change; different readability measure provides either a little bit higher or lower score so in this case, we cannot get any definite trend for this type of comparison. But for each of the assignments, the experiment section

TABLE VI: Lexical Diversity of the Text of the Assignments (Spring 2017)

	Types of Attack		Attack Case Study	
	Control	Expt.	Control	Expt.
Avg. Type-Token Ratio [17]	9.83	15.95	11.92	14.42
Mean Segmental Type-Token Ratio (MSTTR) [18]	5.18	6.40	6.60	7.76
Mean Moving-Average Type-Token Ratio (MATTR) [19]	5.72	7.15	7.90	7.61
Mean Herdan's Constant C (Herdan's C) [20]	3.89	3.77	3.50	3.26
Guiraud's Root Type Token Ratio (Root TTR) [21]	17.97	86.52	51.45	108.23
Carroll's Corrected Type Token Ratio (CTTR) [22]	17.37	76.02	50.85	100.50
Uber Index [23]	19.00	93.92	54.62	133.42
Summer's S	3.27	4.00	3.37	3.95
Yule's K [24]	19.00	94.00	57.5	136.5
Maas' Indices - a [25]	2.86	3.40	3.05	3.26
Maas' Indices - lgV0 [25]	12.35	50.25	30.30	59.11
VocD [26]	19.00	75.15	49.7	81.95
Measure of Textual Lexical Diversity (MTLD) [15]	19.00	96.50	57.10	137.50
Moving-Average Measure of Textual Lexical Diversity (MTLD-MA) [15]	19.00	96.35	57.50	135.23

consistently scores a bit higher. The experiment section has 2.4% on average (of all the measures listed in Table VII) higher readability score for *Types of Cyber Attack Assignment* and 4.3% higher readability for *Cyber Attack Case Study Assignment*. While these differences are not as high as the ones in Section III-C (e.g. VocD or MLTD score), this still shows that the experiment section is approaching their assignments in a different manner than the students from the control section.

E. Hyperlink Usage

We use the number of hyperlinks cited in each assignment as a measure of how the students are looking at outside sources while writing their assignment. The experiment section uses 22% less on-line references for the *Types of Attack* assignment but for the *Attack Case Study* assignment the average usage is same, the variance in the usage among the class is always lower for the experiment section. The results are shown in Figure 13. Although usage of external hyperlinks was encouraged, we feel that since the experiment section attended the presentations and got in-depth info on the types of attack and case studies readily, they did not have to look up information online and cite them, hence the lower number of external references for the experiment section.

F. Objectivity Analysis

In the final stage of our analysis, we apply lexicon based sentiment and objectivity analysis. We want to find out

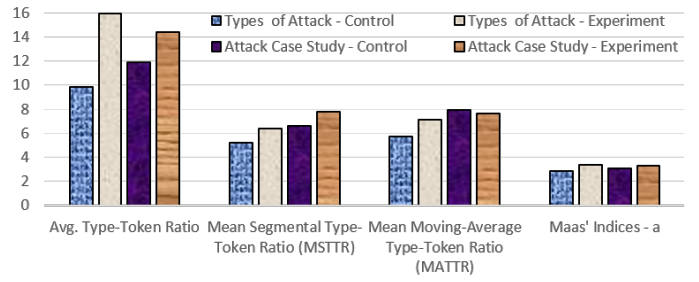


Fig. 7: Comparison of Avg. Type-Token Ratio, MSTTR, MATTR and Maas' Index (a) values of the Text Assignments (Spring 2017)

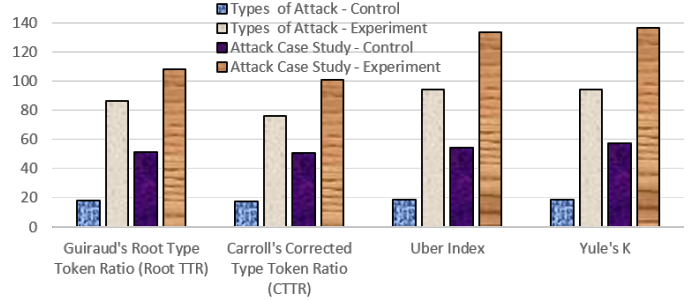


Fig. 8: Comparison of Root TTR, CTTR, Uber Index and Yule's K values of the Text Assignments (Spring 2017)

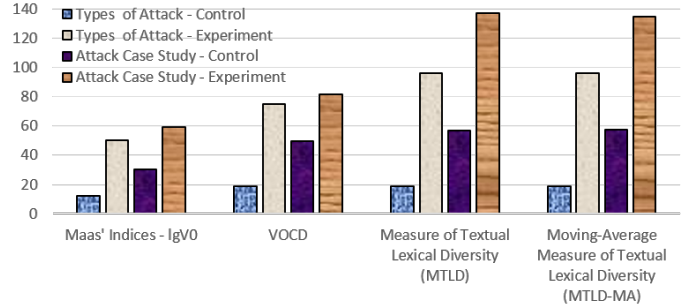


Fig. 9: Comparison of Maas's Index (lgV0), VOCD, MLTD and MLTD-MA values of the Text Assignments (Spring 2017)

whether the peer mentoring sessions are making the students more opinionated as the semester progresses. In the domain of Natural Language Processing, objectivity/subjectivity classification is considered a harder problem than polarity classification [39]. Objectivity and sentiment classification can be done at both document level and sentence level [40]. The former is the simplest form of sentiment analysis which assumes the document contains an opinion on a single context. Sentence level analysis offers more fined grained view by finding the objectivity score at the sentence level and finally expressing the document level sentiment as an aggregation of those scores. Both analyses can be done in either supervised, semi-supervised or unsupervised manner. Lexicon based approaches are a popular unsupervised approaches as there is no inherent need for annotated ground truth data. They generally produce output with high precision but low recall [41]. So for optimal classification, hybrid approaches are often preferred. For our analysis, we went with the lexicon-based approach as we

TABLE VII: Comparison of Readability Measures of the Text of the Assignments Between the Control and Experiment Section (Spring 2017)

	Types of Attack		Attack Case Study	
	Control	Expt.	Control	Expt.
Automated Readability Index (ARI) [30]	9.69	9.81	9.67	10.12
Danielson-Bryan [31]	7.70	7.68	7.65	7.81
Easy Listening Formula (ELF) [32]	5.66	6.01	5.88	6.05
Flesch-Kincaid Grade Level [29]	9.23	9.46	9.30	9.70
Gunning FOG Index [33]	11.66	11.97	11.80	12.15
FORCAST [34]	10.69	10.83	10.50	10.99
Neue Wiener Sachtextformeln(nWS) [35]	6.01	6.17	5.57	6.25
Anderson's Readability Index (RIX) [36]	4.16	4.31	3.96	4.17
Simple Measure of Gobbledygook (SMOG) [27]	11.39	11.53	11.49	11.76
Tuldava's Index [37]	4.30	4.34	4.34	4.43
Läsbarhetsindex (LIX) [36]	43.13	43.66	40.70	42.74
Wheeler-Smith [38]	56.66	60.10	58.81	60.52

lacked annotated ground truth data for supervised analysis. We calculated both document and sentence level objectivity. Document level objectivity was calculated by using *AFINN Word Lexicon* which associates a score between -5 and 5 (from negative to positive sentiment) for each words. Scores between (-1, 1) were assumed to be neutral and anything else was treated as subjective [42]. For each document, these word level scores are aggregated to get the final score. Sentence level objectivity is calculated according to [43] with *MPQA Lexicon* with a similar numeric conversion and aggregation scheme. As shown in Figure 14, the experiment section shows less objectivity for the first assignment (*Types of Attack*) but for the second assignment (*Attack Case Study*) they show more objectivity.

IV. CONCLUSION

In this paper, we have outlined a text mining approach to assess the effectiveness of peer mentoring across two courses in information systems and data analytics for cybersecurity by evaluating the quality of assignments submitted in the peer mentored class vs. class that did not receive the intervention. Throughout the analysis, we have repeatedly observed that *the students from the experiment section who receive peer mentoring are able to use fewer words to express more complex ideas with better readability, lexical diversity and greater objectivity over the span of the semester*. As a result their grades considerably improve which serves to re-validate the findings of peer mentoring in other areas of science, as an effective method of grade improvement. Although in our previous study [12] we were able to compare the results of

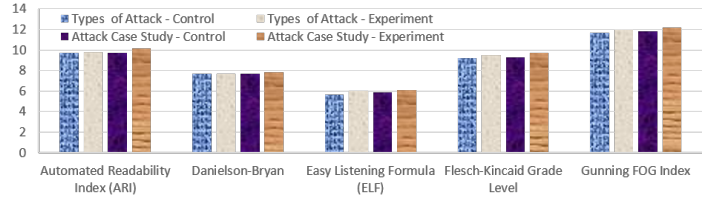


Fig. 10: Comparison of ARI, Danielson-Bryan, ELF, Flesch-Kincaid Grade Level and Gunning FOG Index values of the Text Assignments (Spring 2017)

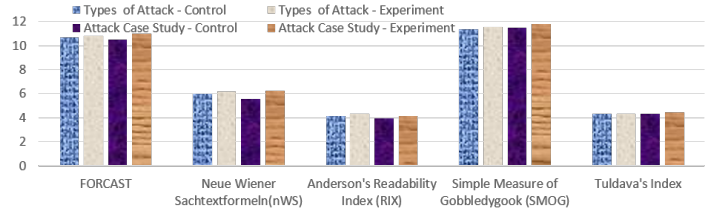


Fig. 11: Comparison of FORCAST, nWS, RIX, SMOG and Tuldava's Index values of the Text Assignments (Spring 2017)

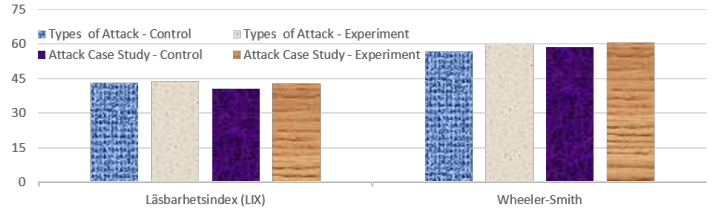


Fig. 12: Comparison of LIX and Wheeler-Smith values of the Text Assignments (Spring 2017)

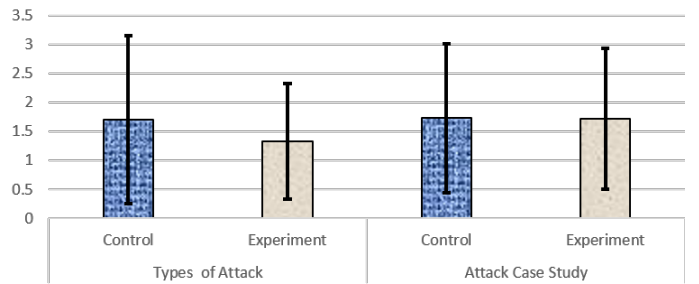


Fig. 13: Average Usage of Hyper-links in 2017 Assignments, (black bars show one standard deviation range)

two semesters of peer mentoring, we only started collecting assignments online from spring 2017 session, so the analysis in this paper is limited to just that semester. We plan to follow up with a detailed comparison with multiple years' result. It would also be interesting to perform a gender-focused analysis (similar to [12]) to investigate whether there is a noticeable difference in the text analysis scores of the male and female students separately between the control and experiment section. Finally, we would like to experiment with more robust supervised and semi-supervised objectivity algorithms for the objectivity classification.

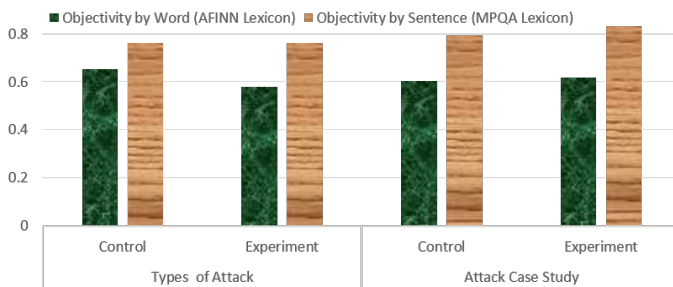


Fig. 14: Objectivity of the 2017 Assignments

V. ACKNOWLEDGMENTS

This work is funded by the National Science Foundation Grant#1515358 titled ‘Integrating Cybersecurity with Undergraduate IT Programs’. The authors would like to thank Dr. Aryya Gangopadhyay, Dr. Carolyn Seaman and Amy Everhart for their insights during the peer mentoring study.

REFERENCES

- [1] J. W. Colvin and M. Ashman, “Roles, risks, and benefits of peer mentoring relationships in higher education,” *Mentoring & Tutoring: Partnership in Learning*, vol. 18, no. 2, pp. 121–134, 2010.
- [2] F. Sana, M. V. Pachai, and J. Kim, “Training undergraduate teaching assistants in a peer mentor course,” *Transformative Dialogues: Teaching & Learning Journal*, vol. 4, no. 3, 2011.
- [3] R. Kirkham and D. Ringelstein, “Student peer assisted mentoring (spam): A conceptual framework-journal of business education & scholarship of teaching,” *E-Journal of Business Education & Scholarship of Teaching*, vol. 2, no. 2, pp. 39–49.
- [4] G. D. Kuh, J. Kinzie, J. H. Schuh, E. J. Whitt, and A. Associates, *Student Success in College: Creating Conditions That Matter*, 1st ed. Jossey-Bass, Mar. 2005.
- [5] A. Kahveci, S. Southerland, and P. Gilmer, “Retaining undergraduate women in science, mathematics, and engineering,” *Journal of College Science Teaching*, vol. 36, no. 3, pp. 34–38, 2006.
- [6] G. Salinitri, “The effects of formal mentoring on the retention rates for first-year, low-achieving students,” *Canadian Journal of Education*, vol. 28, no. 4, pp. 853–873, 2005.
- [7] M. Sorrentino, Diane, “The seek mentoring program: An application of the goal-setting theory,” *Journal of College Student Retention*, vol. 8, no. 2, pp. 241–250, 2007.
- [8] J. P. Dugan and S. R. Komives, “Developing leadership capacity in college students: Findings from a national study,” National Clearinghouse of Leadership Programs, Tech. Rep., 2007.
- [9] S. R. Komives, F. C. Mainella, S. D. Longerbeam, L. Osteen, and J. E. Owen, “A leadership identity development model: Applications from a grounded theory,” *Journal of College Student Development*, vol. 47, pp. 401–418, 2006.
- [10] E. T. Pascarella and P. T. Terenzini, *How College Affects Students : A Third Decade of Research*. Jossey-Bass, Feb. 2005.
- [11] V. P. Janeja, C. Seaman, K. Kephart, A. Gangopadhyay, and A. Everhart, “Cybersecurity workforce development: A peer mentoring approach,” in *Intelligence and Security Informatics (ISI), 2016 IEEE Conference on*. IEEE, 2016, pp. 267–272.
- [12] V. P. Janeja, A. Z. M. Faridee, A. Gangopadhyay, C. Seaman, and A. Everhart, “Enhancing interest in cybersecurity careers: A peer mentoring perspective,” in *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, ser. SIGCSE ’18. New York, NY, USA: ACM, 2018, pp. 384–389. [Online]. Available: <http://doi.acm.org/10.1145/3159450.3159563>
- [13] V. Johansson, “Lexical diversity and lexical density in speech and writing: a developmental perspective,” *Working Papers in Linguistics*, vol. 53, pp. 61–79, 2009.
- [14] D. Malvern, B. J. Richards, N. Chipere, and P. Purán, *Lexical diversity and language development*. Springer, 2004.
- [15] P. M. McCarthy and S. Jarvis, “Mtd, vocd-d, and hd-d: A validation study of sophisticated approaches to lexical diversity assessment,” *Behavior research methods*, vol. 42, no. 2, pp. 381–392, 2010.
- [16] M. Michalke, “korpus: An r package for text analysis,” *Version 0.04-40*, last verified, vol. 11, p. 2013, 2014.
- [17] B. Richards, “Type/token ratios: What do they really tell us?” *Journal of child language*, vol. 14, no. 2, pp. 201–209, 1987.
- [18] R. V. Watkins, D. J. Kelly, H. M. Harbers, and W. Hollis, “Measuring children’s lexical diversity: Differentiating typical and impaired language learners,” *Journal of Speech, Language, and Hearing Research*, vol. 38, no. 6, pp. 1349–1355, 1995.
- [19] M. A. Covington and J. D. McFall, “Cutting the gordian knot: The moving-average type-token ratio (mattr),” *Journal of quantitative linguistics*, vol. 17, no. 2, pp. 94–100, 2010.
- [20] G. Herdan, “Quantitative linguistics,” 1964.
- [21] P. Guiraud, *Problèmes et méthodes de la statistique linguistique*. Presses universitaires de France, 1960.
- [22] J. B. Carroll, “Language and thought,” 1964.
- [23] S. Jarvis, “Short texts, best-fitting curves and new measures of lexical diversity,” *Language Testing*, vol. 19, no. 1, pp. 57–84, 2002.
- [24] C. U. Yule, *The statistical study of literary vocabulary*. Cambridge University Press, 2014.
- [25] H.-D. Mass, “Über den zusammenhang zwischen wortschatzumfang und länge eines textes,” *Zeitschrift für Literaturwissenschaft und Linguistik*, vol. 2, no. 8, p. 73, 1972.
- [26] P. M. McCarthy and S. Jarvis, “vocrd: A theoretical and empirical evaluation,” *Language Testing*, vol. 24, no. 4, pp. 459–488, 2007.
- [27] G. H. Mc Laughlin, “Smog grading-a new readability formula,” *Journal of reading*, vol. 12, no. 8, pp. 639–646, 1969.
- [28] E. Dale and J. S. Chall, “The concept of readability,” *Elementary English*, vol. 26, no. 1, pp. 19–26, 1949.
- [29] J. P. Kincaid, R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom, “Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel,” Naval Technical Training Command Millington TN Research Branch, Tech. Rep., 1975.
- [30] E. A. Smith and R. Senter, “Automated readability index,” *AMRL-TR. Aerospace Medical Research Laboratories (US)*, pp. 1–14, 1967.
- [31] W. A. Danielson and S. D. Bryan, “Computer automation of two readability formulas,” *Journalism Quarterly*, vol. 40, no. 2, pp. 201–206, 1963.
- [32] I. E. Fang, “The “easy listening formula,”” *Journal of Broadcasting & Electronic Media*, vol. 11, no. 1, pp. 63–68, 1966.
- [33] R. D. Powers, W. A. Sumner, and B. E. Kears, “A recalculation of four adult readability formulas,” *Journal of Educational Psychology*, vol. 49, no. 2, p. 99, 1958.
- [34] G. R. Klare, “Assessing readability,” *Reading research quarterly*, pp. 62–102, 1974.
- [35] R. Bamberger and Vanacek, *Lesen-Verstehen-Lernen-Schreiben*. Diesterweg, 1984.
- [36] J. Anderson, “Lix and rix: Variations on a little-known readability index,” *Journal of Reading*, vol. 26, no. 6, pp. 490–496, 1983.
- [37] J. Tuldava, “Measuring text difficulty,” *Glottometrica*, vol. 14, pp. 69–81, 1993.
- [38] L. R. Wheeler and E. H. Smith, “A practical readability formula for the classroom teacher in the primary grades,” *Elementary English*, vol. 31, no. 7, pp. 397–399, 1954.
- [39] B. Pang, L. Lee *et al.*, “Opinion mining and sentiment analysis,” *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, pp. 1–135, 2008.
- [40] R. Feldman, “Techniques and applications for sentiment analysis,” *Communications of the ACM*, vol. 56, no. 4, pp. 82–89, 2013.
- [41] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, *Combining lexicon-based and learning-based methods for twitter sentiment analysis*, 89th ed., 6 2011.
- [42] L. K. Hansen, A. Arvidsson, F. Å. Nielsen, E. Colleoni, and M. Etter, “Good friends, bad news-affect and virality in twitter,” in *Future information technology*. Springer, 2011, pp. 34–43.
- [43] T. Wilson, J. Wiebe, and P. Hoffmann, “Recognizing contextual polarity in phrase-level sentiment analysis,” in *Proceedings of the conference on human language technology and empirical methods in natural language processing*. Association for Computational Linguistics, 2005, pp. 347–354.