Scientific Research

http://dx.doi.org/10.4236/sm.2013.32026

Study of Pauline Epistles in the New Testament Using Machine Learning

Wei Hu

Department of Computer Science, Houghton College, New York, USA
Email: wei.hu@houghton.edu

The New Testament of the Bible is a collection of 27 books, 13 of which are attributed to St. Paul, making him a significant author of the New Testament. The Pauline letters were written to churches and individuals addressing various issues in the early church and offering instructions and encouragement to the believers of his day. To understand Paul's message as a whole, we clustered his letters, their chapters, and their verses respectively to identify the topics that are addressed repeatedly or once, offering different views at three granularities with each complementing the other. Our calculation implied all Pauline letters could be divided into six groups, well in line with the church tradition. They were First and Second Thessalonians, First and Second Corinthians, Colossians and Ephesians, Galatians and Romans, Philemon and Philippians, and First and Second Timothy and Titus. Among all letters First Corinthians had the most unique chapters, while First Corinthians and Romans had the most unique verses. Furthermore, similar verses and chapters from his letters were also discovered. We also studied the authorship of the letter to Hebrews, an anonymous letter in the New Testament. Using known Pauline and non-Pauline letters as training data, we built nine highly reliable predictors, which collectively predicted that the book of Hebrews was not authored by Paul.

Keywords: Bible; New Testament; Paul; Topic Model; Clustering; Affinity Propagation

Introduction

The Bible is a collection of books written by human authors inspired by God, in which God progressively reveals Himself to man in creation, in history, especially in the history of Israel, and ultimately in His son Jesus Christ. It is made of the Old Testament comprising 39 books and the New Testament containing 27 books. The books in the Old Testament could be categorized into books of the law, history books, wisdom books, and books of the prophets, while those in the New Testament can be divided into the Gospels, Acts of the Apostles, epistles, and revelation. The four Gospels describe the life and ministry of Jesus, the Acts record the early days of the church, the epistles are 21 books written in the form of letters to churches or individuals, and the book of Revelation presents the revelation of Jesus Christ. Of the 21 letters in the New Testament, St. Paul wrote 13 of them. Some scholars regard the anonymous epistle to Hebrews a 14th Pauline letter. The letters of Paul, addressed to a given individual or community of his day, are the earliest existing Christian writings, predating the four gospel books in the New Testament. They represent a particular form of early Christianity, which clearly is different from that developed in Jerusalem. The former is Gentile and the latter is Jewish.

The Apostle Paul was born, as a Roman citizen, in the city of Tarsus in the province of Cilicia (southern Turkey today) with a Hebrew name Saul given at birth and a Latin name Paul. At young age, he went to Jerusalem to study under Gamaliel, one of the greatest rabbinical teachers of the first century (Acts 5:24). He was circumcised on the eighth day, of the people of Israel, of the tribe of Benjamin, a Hebrew of Hebrews; in regard to the law, a Pharisee; as for zeal, persecuting the church; as for legalistic righteousness, faultless (Philippians 3:5-6).

One day on his way from Jerusalem to Syrian Damascus on a mission to persecute followers of Jesus, Paul encountered and conversed with the resurrected Jesus. As a result he was converted from a zealous Pharisee who aggressively persecuted the Christians to a fervent preacher of the gospel to the Gentiles, a radical change in his life made by the saving grace of God. This conversion influenced Paul's viewpoint from "righteousness based on the law" to "righteousness based on the death of Christ alone", which later became the centerpiece of his theology. There was also a shift of referring to him as Saul to Paul as recorded in Acts 13:9. It is speculated that the change to use his Gentile name might be an indicator of his mission as an Apostle to Gentiles. In the book of Acts, Luke devotes the majority of the book to the life and journeys of Paul, including Paul's three missionary journeys to spread the news of salvation through the death of Jesus. As a result he established several churches in different regions including Ephesus, Corinth, Galatia, Thessalonica, and Philippi. However, the churches in Roman and Colossae were not established by Paul. Well-versed in Jewish theology and language, and the Greek language and culture, Saul was God's chosen instrument to deliver the gospel to the Gentiles, as the message from Christ to Paul through Ananias in Acts says: "Go! This man is my chosen instrument to carry my

name before the Gentiles and their kings and before the people of Israel" (Acts 9:15).

In his work to create and strengthen the early churches, Paul devoted much of his attention to writing letters to the churches and individuals to address some relevant issues or topics. These letters were an effective means of instructing a community from a distance, since during his time travel and communication were slow and difficult. His letters aimed to help the corresponding congregations to be firmly rooted in the gospel of Christ, which then naturally lead to his call for gospel-centered living. The perfect balance between doctrine and the practice of the doctrine is characteristic of his letters (Kostenberger & Wilder, 2010).

Of all Paul's writings, 13 of them are collected in the New Testament and constitute a critical library of books that churches and Christians use today. Unfolding the teachings of Christ, the Pauline letters develop the doctrine of Christian faith and the church including salvation, justification, reconciliation, sanctification, and glorification. Paul's writings are both letter and epistle, since they often retain personal greetings, recommendations of people and at the same time contain teachings. Jesus used a simple language to describe the kingdom of God, which was full of comparisons and parables taken from daily life. In contrast, Paul typically used a more complex language than that of Jesus to those people who had heard the gospel. The letters of Paul have touched almost every important doctrine held by the Christian church today, emphasizing the strong connection between faith and faith-based living as the fruits of the spirit. Therefore they are really worthy of attentive and earnest study. As Psalm 119:105 says, God's word is a lamp to our feet and a light for our path.

The aim of our work was to study all Paul's writings through clustering his letters, their chapters, and their verses to categorize all topics covered by Paul. With this information, we were able to develop a global view at different resolutions of the important writings of Paul in the New Testament of the Bible.

Materials and Methods

The Text of the Bible

The text of the Bible used in this study is from the King James version (1611 authorized version), downloaded from http://printkjv.ifbweb.com. Because the letters of Paul were used as data in our study, this section provided some background information about them.

The Pauline Letters

The letters of Paul could be classified into three groups in their chronological order (**Table 1**): travel letters (First & Second Thessalonians, Galatians, First & Second Corinthians, and Romans), prison letters (Colossians, Ephesians, Philemon, and Philippians), pastoral letters (First & Second Timothy, and Titus). His letters tend to be written in response to specific issues, and nine of them are addressed to churches while four to individuals. These letters are known by the names of their recipients. Their main purpose is to teach, encourage, and strengthen the small house churches he was writing to, helping them collectively to lay their foundation on the gospel of Christ. Also the individual members in these churches were in urgent need of further enlightenment on daily Christian life, which reflects the very practical aspect of Paul's writings. To this end, the letters of Paul demonstrate a close connection between faith and living that is a clear trait of Paul's writings. Below we will briefly summarize the contents of these letters.

Travel Letters

Paul's missionary work in Thessalonica is described in Acts 17. The key topic in the First & Second Thessalonians is the return of the Lord, especially its timing. Paul explains the unexpectedness of Christ's return and warns the believers for a lifelong preparedness. The second letter to Thessalonians attempts to correct the misunderstanding of the second coming of the Lord, a topic Paul brings up in his first letter to Thessalonians.

The churches in Galatia were created by Paul on his first missionary journal. The Gentile believers in Galatia were told by Jewish believers that they should follow the Jewish law along with faith in Christ, implying keeping the Law of Moses is essential to their salvation. In this letter to Galatians, he defends his position as an apostle, reaffirms the doctrine of justification by faith alone, and explains the freedom that the believers enjoy from the total sufficiency of faith in Christ, highlighting the redemptive death of Jesus on the cross. The word "justify" means to "declare righteous." The issue of righteousness has been a historical one. The book of Job in Old Testament asks, "How can a mortal be righteous before God?" (Job 9:2), and the book of Habakkuk declares, "The righteous will live by his faith" (Habakkuk 2:4). In this letter, Paul describes faith as the sole means by which people can attain salvation and as the only source of eternal life. This grace of God gives the believers the freedom and power not to sin and to love one another, as the famous verse in this letter says: "The fruit of the Spirit is love, joy, peace, patience, kindness, goodness, faithfulness, gentleness, and self-control" (Galatians 5:22). It teaches us how to relate to ourselves, others, and God. In this verse, Paul makes a strong connection between faith and holiness expressed as the fruits of the faith.

Table 1.
Timeline by year of Paul's letters (http://www.blueletterbible.org/study/pnt/pnt02.cfm).

A Basic Time Line of Paul's Letters										
Conversion of Paul	First Missionary Journey Begun	Secondary Missionary Journey Begun		Third Missionary Journey Begun			First Imprisonment at Rome		Second Imprisonment at Rome	
36	48	51	52 - 53	54	57	58	62	63	65	66
			Thessalonians		Galatians 1Corinthians	2Corinthians Romans	The Prison Epistles	1Timothy, Titus		2Timothy

The letter to Romans is the longest and regarded as most important letter of Paul. It addresses many issues that are fundamental to Christian faith and its practice, an apparent continuation and elaboration of Paul's theology stated in Galatians. The theological principles covered in this letter include condemnation, justification, sanctification, glorification, and consecration, with a distinct exposition on salvation by the grace of God through faith in Jesus Christ alone. In this letter, Paul introduces himself to the believers of the Roman church and strengthens their faith, and informs them his plan to visit their city and hope they will assist him financially go from there to Spain (Romans 15:24,28). He plans to use Rome as a base for his missionary work similar to that in Antioch, since Spain is far from Antioch. He also deals with some of the problems that the church in Rome is dealing with such as the unity between Jewish and Gentiles believers.

He begins in this letter with an explanation of the sinful state of humanity that makes it impossible to keep the law in the sight of God. He writes "For all have sinned, and come short of the glory of God." (Romans 3:23). He then teaches that salvation from sin is only possible by the grace of God through faith in Jesus and not through observance of Jewish law using Abraham as an example. He writes: "For the wages of sin is death; but the gift of God is eternal life through Jesus Christ our Lord" (Romans 6:23). He further points out that "A man is a Jew if he is one inwardly; and circumcision is circumcision of the heart, by the Spirit, not by the written code" (Romans 2:29). Paul also advises the Romans to live as a follower of Jesus with harmony, humility, love, charity, forgiveness, and submission, since he is aware of some conflict between Gentile and Jewish believers in the Roman church. Paul presents his doctrine of justification through faith in Galatians and Romans based on the examination of the Mosaic Law and quoting the story of Abraham's faith. In these two letters, the gospel of grace, in contrast to the false gospel of works, is revealed more clearly than in his other letters.

The major theme (or a concise summary of the content) of this letter is "It is the power of God for the salvation of everyone who believes; first for the Jew, then for the Gentile. For in the Gospel, a righteousness from God is revealed, a righteousness that is by faith from first to last, just as it is written: "he righteous will live by faith" (Romans 1:16-17). Here Paul quotes a verse in Habakkuk 2:4: "The righteous will live by faith", a truth illustrated throughout the history of Israel and first modeled in Abraham's life. He also quotes this verse in Galatians 3:11. Further, Hebrews 10:38 uses this verse as well. The death of Jesus on the cross reconciled the sinner, Jews and Gentiles, with God, which signals the end of Law as a means of justification. Paul emphasizes justification is by the grace of God through faith in Christ and the children of God are not based on genealogy or circumcision, but on faith in Jesus (Stott, 2001).

First Corinthians is written to correct many problems developed in the Corinthian community (divisions in the church, trust in man's wisdom, immorality) and to answer the questions from the church regarding marriage and divorce, food, worship, spiritual gifts, and resurrection. His own conversion provides powerful testimony to the resurrection of Christ. This letter also contains the greatest description and explanation of love ever written (1 Corinthians 13). In regard to the wisdom of God, Paul writes "The foolishness of God is wiser than man's wisdom, and the weakness of God is stronger than man's strength"

(1 Corinthians 1:25).

After the first letter to Corinthians, news of continuing problems in Corinth reached Paul. The second letter to Corinthians is the least doctrinal of Pauline letters, but an intensely personal and emotional letter revealing the personal character of the apostle more than any other letters. At the same time, the content structure of this letter is evidently uneven and digressive, as seen from the sudden shift of subject between chapters 10 - 13 and chapters 1 - 9. As a result, some scholars suggest that this letter is a composition of a few letters of Paul. In this letter, Paul affirms his affection for the Corinthians and expresses his joy at the favorable response of the church to his ministry. But most importantly, he defends the authenticity of both his apostleship and his message. His final greetings in chapter 13:14 ("May the grace of the Lord Jesus Christ, and the love of God, and the fellowship of the Holy Spirit be with you all.") are frequently used in the church today.

Prison Letters

After hearing from a church leader Epaphras (Colossians 1:7-8) about the gospel had born fruit and some dangers (paganism and heresy) facing the Colossian church, Paul wrote Colossians to encourage the church in their faith in Christ and elaborates the true gospel based on the supremacy and sufficiency of Christ, and to guard them against the dangers facing the church. Paul claims as in Ephesians that Christ is the head of the church (Colossians 1:18). Colossians focuses on Christ as the header of the church, while Ephesians teaches the concept that church is the body of Christ.

Ephesians was written after Colossians, dealing with the similar subjects in a more elaborate way and having a style closer to a doctrinal letter (O'Brien, 1999; Goodspeed, 1933). Colossians has 4 chapters, but Ephesians has 6. It does not deal with problems of particular churches; instead it points the views of its readers from an earthly/temporary reference to a heavenly/eternal reference. In this letter, Paul reveals the mystery that Christ is the head of the church and the church is the body of Christ. To keep the church pure and holy as Christ is pure and holy, he encourages the believers to live in a manner that is fitting with their position as believers. It contains no allusion to the personal circumstances of people in the city of Ephesus. The letter to Ephesians can be divided into two parts. The first part, chapters 1 - 3, explains the doctrine, while the second part, chapters 4 - 6, deals with the practice of Christian faith. He writes: "For it is by grace you have been saved, through faith—and this not from yourselves, it is the gift of God—not by works, so that no one can boast" (Ephesians 2:8-9). The Ephesians describes the church as the body of Christ and God's glory in the church, while in Corinthians Paul discusses the details of establishing the church as a witness to Christ on earth.

Paul established at Philippi the first Christian community in Europe on his second missionary journey (Acts 16:9-40). Comparing to his other letters, Philippians could be read as a letter of joy, unity, and friendship in Christ. He states explicitly his love for them such as "I have you in my heart" and "I long for you with the affection of Christ Jesus" (Philippians 1:7-8) and his thanksgiving and acknowledgement of their gifts such as "I always pray with joy because of your partnership in the gospel from the first day until now" (Philippians 1:4) and "you sent me aid again and again when I was in need (Philippians

4:16). These verses clearly suggest a strong mutual affection between Paul and his converts in Philippi, not found in any other churches. He does not open this letter with his credential as an apostle, since he does not need to do so. He also informs the church of his imprisonment situation and warns them against the false teachers.

Philemon was a leader in the church at Colossae. One of his slaves Onesimus had run away and found Paul imprisoned in Rome, and became a convert and very useful to Paul. In his letter to Philemon, Paul appeals to Philemon to forgive Onesimus and receive him as a brother in Christ.

Pastoral Letters

The greatest need in the church is effective leadership as shown in the church history. Three pastoral letters were written to Timothy and Titus, two young church leaders and close co-workers of Paul. These letters resemble one another in their content, and are exhortations and practical instructions on the life, doctrine, and ministry of the church including qualifications for church leaders, church organization, worship, and ethics in an effort to create a highly institutionalized church. First Timothy and Titus share many topics. Like the second letter to Corinthians, Second Timothy is a very personal letter, containing the farewell remarks from Paul to a young coworker.

Affinity Propagation

The affinity propagation (AP) clustering algorithm (Frey & Dueck, 2007) aims to search exemplars from data to represent clusters through message passing. The data points within a cluster connect to the exemplar that best represents them.

Topic Model

Topic models explore the correlation between the observed words in documents and the latent topics in the documents, where each topic is defined by a group of related keywords (Steyvers & Griffiths, 2007; Griffiths, Steyvers, & Tenenbaum, 2007; Griffiths & Steyvers, 2004). The Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) is a simple topic model that assumes each document is a mixture of various topics. It uses generative probabilistic model to represent the topics in documents as probabilistic distributions, which could then be used to estimate the similarity between documents.

In this study, the topic distribution identified by LDA from each document is used to define a correlation to measure the similarity between two documents. The standard Pearson correlation formula for two vectors in n-dimensional Euclidean space is defined as follows:

$$r_1 = \frac{\sum_{i=1}^n (\bar{X} - X_i)(\bar{Y} - Y_i)}{(n-1)S_X S_Y}$$

where $X = (X_1, X_2, X_3, \cdots X_n)$, $Y = (Y_1, Y_2, Y_3, \cdots Y_n)$. $\bar{X}$ and $\bar{Y}$ represent the average of the two vectors, S_X and S_Y stand for the standard deviation of the two vectors.

In the calculation of the correlation between two topic distributions from two documents, the X_i represents the probability of topic i, and the sum of all these X_i is one. Since the contributions from different X_i to the correlation are not equal, we assign a weight to each expression $(\bar{X} - X_i)$ and $(\bar{Y} - Y_i)$ as shown below:

$$r_2 = \frac{\sum_{i=1}^n X_i (\bar{X} - X_i)(\bar{Y} - Y_i) Y_i}{(n-1)S_X S_Y}$$

which suggests that the higher the probability of a topic is, the greater its contribution to the correlation. Throughout the current study, we used this weighted Pearson correlation r_2 for comparing two topic distributions.

In the present study, the AP clustering algorithm was used to group the Pauline letters, their chapters, and their verses, where each letter/chapter/verse segment was represented by its topic distribution as a vector in Euclidean space.

Results

In this section, we report the outcome of out computational analysis of Paul's letters, conducted at three different resolutions: letters, chapters, and verses. Our aim was to apply the topic modeling and AP clustering algorithm to catalog the topics discussed either once or repeatedly in all his writings, thus rendering the whole picture of his messages.

Clusters of 13 Pauline Letters

We first clustered the 13 letters of Paul, which produced five groups (**Table 2**). Because of the different contents of First and Second Corinthians, they were not placed inside the same cluster. But these two letters had many similar chapters as seen in "Clusters of All Chapters of Pauline Letters" Section. All the other Pauline letters were correctly grouped according to the church tradition. Although there are 13 letters, but our clustering algorithm was able to find five exemplars from them: Colossians, Philippians, Romans, First Thessalonians, and First Timothy.

The word clouds in **Figure 1** gives the visual display of the most frequently occurring words in Paul's letters. In Romans and Galatians, the words "law", "faith", "righteousness", "spirit", "sin", "flesh" are emphasized. In pastoral letters, the word "good" stands out; it is used to describe "good work(s)", "good conscience". This word is mainly used in First Timothy and Titus, but not in Second Timothy. In First and Second Thessalonians, the word "brethren" (meaning brothers) appears frequently, showing Paul's personal affection towards the believers in that church. The word "Christ" is mentioned more often than the word "God" in the groups of Colossians and Ephesians, and Philippians and Philemon than in other groups.

Pairwise Correlations of Pauline Letters

Here we looked at the correlations between a pair of the Pauline letters (**Table 3**), which was not readily available through the clustering of these letters in "Clusters of 13 Pauline Letters" Section. It turned out the most similar letters within the Pauline letters were Colossians and Ephesians, Philemon and Philippians, First and Second Thessalonians, First Timothy and Titus, and Galatians and Romans (**Table 3**). This pairwise comparison refined the clusters in **Table 2**. Say, the three pastoral letters were in the same group in **Table 2**, but actually the First Timothy and Titus were the most similar of the three, which is true when we compare the contents of these three letters. Moreover, First Corinthians deals with many specific issues at the church, so it is not strongly related to other letters of Paul and is what makes it stand out. We should point out that the thorough pairwise comparison of these letters was only

Table 2.
Clusters of the Pauline letters.

Cluster Number	Exemplar Letter Name	Letters in One Cluster
1	Colossians	Colossians, Corinthians1, Ephesians
2	Philippians	Corinthians2, Philemon, Philippians
3	Romans	Galatians, Romans
4	Thessalonians1	Thessalonians1, Thessalonians2
5	Timothy1	Timothy1, Timothy2, Titus

Table 3.
Top 48 highly correlated pair of Pauline letters.

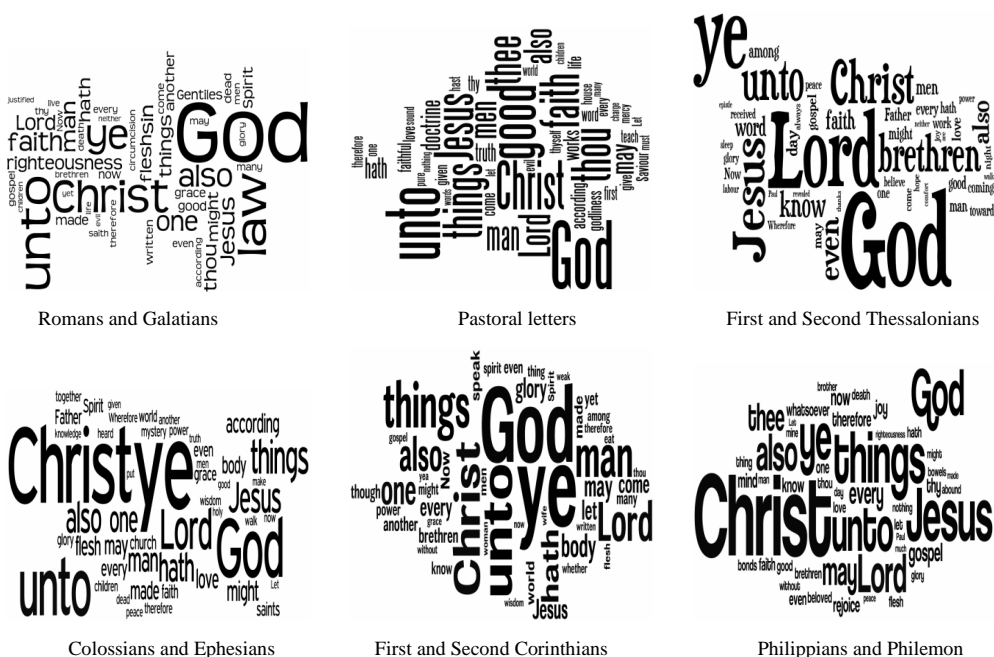| Book1	Book2	Correlation	Book1	Book2	Correlation
Colossians	Ephesians	.983	Corinthians2	Philemon	.829
Philemon	Philippians	.983	Ephesians	Timothy2	.822
Thessalonians1	Thessalonians2	.982	Philippians	Timothy2	.821
Ephesians	Philippians	.976	Ephesians	Galatians	.811
Philippians	Thessalonians2	.96	Philemon	Timothy2	.809
Ephesians	Philemon	.957	Colossians	Timothy2	.808
Timothy1	Titus	.954	Galatians	Philippians	.793
Philippians	Thessalonians1	.952	Colossians	Galatians	.789
Ephesians	Thessalonians2	.95	Timothy1	Timothy2	.789
Colossians	Philippians	.948	Galatians	Thessalonians2	.785
Philemon	Thessalonians2	.944	Galatians	Thessalonians1	.776
Ephesians	Thessalonians1	.937	Galatians	Philemon	.763
Philemon	Thessalonians1	.933	Ephesians	Romans	.749
Colossians	Philemon	.932	Romans	Thessalonians2	.749
Colossians	Thessalonians2	.928	Romans	Thessalonians1	.741
Colossians	Thessalonians1	.918	Colossians	Romans	.737
Galatians	Romans	.888	Corinthians2	Galatians	.724
Corinthians2	Philippians	.862	Philippians	Romans	.721
Corinthians2	Ephesians	.851	Corinthians2	Timothy2	.718
Thessalonians2	Timothy2	.842	Galatians	Timothy2	.695
Thessalonians1	Timothy2	.839	Corinthians2	Romans	.687
Corinthians2	Thessalonians1	.835	Philemon	Romans	.682
Corinthians2	Thessalonians2	.834	Romans	Timothy2	.678
Colossians	Corinthians2	.829	Timothy2	Titus	.673

possible with the help of computational methods.

Clusters of All Chapters of Pauline Letters

There are 87 chapters in the 13 letters of Paul, from which our clustering algorithm generated 21 clusters (**Table 4**). These clusters of various sizes revealed the structural grouping of the different topics covered in the letters. In Colossians, there are 8 verses on rules for Christian households (verses 3:18-25), but in Ephesians, Paul used 21 verses on the same topic (verses 5:22-6:9). As a result, chapter 3 of Colossians and chapter 5 of Ephesians were grouped together in cluster 8.

Paul's letters typically include doctrine and its practice. For instance, Ephesians has doctrine in chapters 1 - 3 grouped in cluster 12, and its practice in chapters 4 - 6, with chapter 4 in cluster 21, chapter 5 in cluster 8, and chapter 6 in cluster 19. Interestingly, these clusters (8, 12, 19, and 21) had the most chapters, and the chapters from different letters in these 4 largest clusters tended to be similar. Further, each of them selected the chapters from specific letters. For example, cluster 8 selected its most chapters from First and Second Corinthians, cluster 19 liked the chapters from Second Corinthians and First and Second Thessalonians, and cluster 21 favored the chapters from Romans, First and Second Timothy. It was noteworthy that First and Second Corinthians were in two different clusters in "Pairwise Correlations of Pauline Letters" Section when

Figure 1.
Word clouds of top 50 frequent words in six groups of Paul's letters.

Table 4.
Clusters of chapters of Pauline letters (To save space, we used abbreviation. Say, Corinthians2_4-6 stands for Second Corinthians chapters 4 - 6.).

Cluster Number	Chapters in one cluster: Subjects of the chapters	Summary of large clusters
1	Corinthians1_11: Propriety in worship, The Lord's supper	
2	Corinthians1_12: Spiritual gifts, One body, many parts	
3	Corinthians1_13: Love	
4	Corinthians1_14: Gifts of prophecy and tongues, Orderly worship	
5	Corinthians1_15: The resurrection of Christ, The resurrection of the dead, The resurrection of the body	Doctrine of the resurrection
6	Colossians_4: Further instructions, final greetings Corinthians1_16: The collection for God's people, Personal Request, Final Greetings Philemon_1: Thanksgiving and prayer, Paul's plea for Onesimus Philippians_4: Exhortations, Thanks for their gifts, final greetings Romans_16: Personal greetings	Final greetings
7	Corinthians1_3: On divisions in the church	
8	Colossians_2-3: Freedom from human regulations through life with Christ, Rules for holy living, Rules for Christian households Corinthians1_2: Wisdom from the spirit Corinthians1_4-6: Apostles of Christ, Expel the immoral brothers, Lawsuits among believers, Sexual immorality Corinthians1_9: The right of Apostles Corinthians2_4-6: Treasures in jars of clay, Our heavenly dwelling, The ministry of reconciliation, Paul's hardship, Do not be yoked with unbelievers Corinthians2_10-12: Paul's defense of his ministry, Paul and the false apostles, Paul's boasts about his sufferings, Paul's vision and his thorn, Paul's concern for the Corinthians Ephesians_5: Living as children of light, Wives and Husbands	Daily Christian living including husband and wife relationship and Paul's exposition on his ministry (mainly from chapters in First and Second Corinthians)
9	Corinthians1_7: Marriage	

Continued

10	Corinthians1_8: Food sacrificed to idols Corinthians1_10: Warnings from Israel's history, Idol feast and Lord's supper, the believer's freedom Romans_14: The weak and the strong	Food related to the church life
11	Corinthians2_3: The glory of the new covenant	
12	Colossians_1: Thanksgiving and prayer, The supremacy in Christ, Paul's labor for the church Corinthians1_1: Thanksgiving, Divisions in the church, Christ the wisdom and power of God Ephesians_1-3: Spiritual blessings in Christ, Thanksgiving and prayer, Made alive in Christ, One in Christ, Paul the preacher to the Gentiles, A prayer for the Ephesians Galatians_1: No other Gospel, Paul called by God Romans_15: Paul the minister to the Gentiles, Paul's plan to visit Rome	Daily Christian living and Paul's exposition on his ministry (First chapters from different letters)
13	Galatians_3: Faith or observation of the law, The law and the promise, Sons of God Romans_4: Abraham justified by faith	Doctrine of Justification: Justification by faith
14	Galatians_4: Paul's concern for the Galatians, Hagar and Sarah	
15	Galatians_5-6: Freedom in Christ, Life by the spirit, Doing good to all, Not circumcision but a new creation Romans_8: Life through spirit, Future glory, More than conquerors	Doctrine of sanctification: Christian's freedom in Christ
16	Romans_11: The remnant of Israel, Ingrafted branches, All Israel will be saved	Doctrine of Election: Election by grace, not by deeds
17	Galatians_2: Paul accepted by the apostles, Paul opposes Peter Romans_2-3: God's righteous judgment, The Jews and the law, God's faithfulness, No one is righteous, Righteousness through faith Romans_9-10: God's sovereign choice, Israel's unbelief	Doctrine of Justification: Righteousness through faith in Christ, not by the deeds of the Law Doctrine of Election: Election by grace, not by deeds
18	Romans_6-7: Dead to sin, alive in Christ; Slaves to righteousness, An illustration from marriage, Struggling with sin	Doctrine of sanctification: Dead to sin, freed from sin, dead to the requirements of the law, struggle with sin
19	Corinthians2_1-2: God of all comfort, Paul's change of plans, Forgiveness for the sinner, Ministers of the new covenant Corinthians2_7-9:Paul's joy, Generosity encouraged, Titus sent to Corinth, Sowing generously Corinthians2_13, Final warnings, Final greetings Ephesians_6, Children and parents, Slaves and masters, The armor of God, Final greetings Philippians_1-3: Thanksgiving and prayer, Paul's chains advance the Gospel, Imitating Christ's humility, Shining as stars, Timothy and Epapharodius, No confidence in the flesh, Pressing on toward the goal Thessalonians1_1-5: Thanksgiving for the Thessalonians' faith, Paul's ministry in Thessalonica, Paul's longing to see the Thessalonians, Timothy's encouraging report, Living to please God, The coming of the Lord, Final instructions Thessalonians2_1-3: Thanksgiving and prayer, The man of lawlessness, Stand firm, Request for prayer, Warning against idleness, Final greetings Timothy2_1: Encouragement to be faithful	Daily Christian living and Paul's exposition on his ministry (First chapters and last chapters from different letters, mainly from chapters in Second Corinthians, Philippians, First and Second Thessalonians)
20	Timothy1_3: Overseers and deacons Timothy1_5: Advice about widows, elders, and slaves Titus_1-2: Titus's task on Crete, What must be taught to various groups	Dealing with various businesses in the church
21	Ephesians_4: Unity in the body of Christ, Living as children of light Romans_1: Paul's longing to visit Roman, God's wrath against mankind Romans_5: Peace and joy; Death through Adam, life through Christ Romans_12-13: Living sacrifice, Love, Submission to the authorities; Love, for the day is near Timothy1_1-2: Warning against false teachers of the law, The Lord's grace to Paul, Instructions on worship Timothy1_4: Instructions to Timothy Timothy1_6: Love of money, Paul's charge to Timothy Timothy2_2-4: A workman approved by God, Godlessness in the last days, Paul's charge to Timothy, Personal remarks, Final greetings Titus_3: Doing what is good, Final remarks	Christian daily living as a transformed life reflected in love, including instructions to Timothy and Titus as individuals (mainly from chapters in Romans, pastoral letters)

compared with their whole letter contents. But many of their chapters were still similar as seen in cluster 8, implying that chapter-wise clustering could reveal new information about these letters. Cluster 12 featured many opening chapters of the letters, including Colossians, First Corinthians, Ephesians, and Galatians. Chapter 15 of Romans was included in this cluster since Paul explains his role as a minister in this chapter, which typically occurs in the opening statements of his other letters.

In total there were 3 clusters, clusters 6, 12, and 19, containing the beginnings and/or ends of the letters. Cluster 6 had the final greetings of 5 letters: Colossians, First Corinthians, Philemon, Philippians, and Romans. Cluster 19 contained both the beginnings and ends of several letters. It had the opening chapters of 5 letters: Second Corinthians, First and Second Thessalonians, Second Timothy, and Philippians. The final chapters in this cluster were from 4 letters: Second Corinthians, Ephesians, First Thessalonians, and Second Thessalonians. As a matter of fact, all chapters of First and Second Thessalonians, and the first 3 out of all 4 chapters of Philippians were in this cluster. These 3 clusters suggested that our clustering algorithm was able to identify the structural features of the Pauline letters, since they do resemble one another in the beginning or ending.

Several similar chapters of Galatians and Romans were grouped together. Clusters 13 had chapter 3 of Galatians and chapter 4 of Romans, both of which mention Abraham as an example of righteousness through faith. Cluster 15 contained chapters 5 and 6 of Galatians and chapter 8 of Romans, dealing with the doctrine of sanctification.

In general, the cluster of a single chapter indicate the topics within the chapter are only discussed once in Paul's letters, while the existence of several chapters in one cluster suggests same or similar topics are discussed in multiple chapters. For example, chapter 3 of Colossians and chapter 5 of Ephesians were in cluster 8, because these two chapters deal with the relationship between husband and wife. The clusters that each contained only a single chapter were mainly made of the chapters from the First Corinthians (chapters 3, 7, 11 - 15) and chapter 3 from the Second Corinthians in cluster 11, chapter 4 from Galatians in cluster 14, and chapter 11 from Romans in cluster 16. Therefore the topics in these chapters of First Corinthians, Galatians, and Romans were unique.

To further visually enlighten the correlations of chapters in the Pauline letters, we chose the top 50 highly correlated pairs of chapters as seen in **Figure 2**, where each pair of chapters were connected with a straight line. The largest connected group in **Figure 2** was made of first chapters of Colossians, Ephesians, First and Second Thessalonians, Philippians, Philemon, and Second Corinthians, and First and Second Timothy. This closely connected network of first chapters enlarged the grouping illustrated in the clusters 12 and 21 in **Table 4**. The 4 letters whose first chapter was not included in this group were First Corinthians, Galatians, Romans, and Titus. There were 5 letters whose final chapters were all in cluster 6 in **Table 4**. Here the final chapters of Romans and First Corinthians were connected as a refinement of the information revealed in cluster 6.

Figure 2.
Top 50 highly correlated pairs of chapters in Paul's letters where each pair was connected with a line.

Clusters of All Verse Segments in the Pauline Letters

Here we divided each chapter into numerous segments of 10 verses with the last segment having 10 verses or less than. As a result, there were 231 segments created from all the 13 letters of Paul. Our clustering algorithm produced 33 clusters with the largest cluster, cluster 23, holding 158 segments, which showed the similarity of the most segments in Paul's letters (**Table 5**). As in the discussion of different chapters in "Clusters of All Chapters of Pauline Letters" Section, it was of interest to look at the topics appeared in the small clusters, some of which overlapped with those in "Clusters of All Chapters of Pauline Letters" Section but some didn't.

In Romans, Paul greets many people by name in the Rome church (Romans 16:1-21). In general, Paul does not offer such extensive greetings in his letters. Our algorithm picked on Paul's personal request and greetings which contain several names in First Corinthians along with Romans 16:1-21 in cluster 24. The two letters, Romans and First Corinthians, had the most unique verses or topics as detected by our algorithm. Cluster 31 contained a verse segment that deals with the second coming of the Lord, a topic not found in the chapter clusters in "Clusters of All Chapters of Pauline Letters" Section, thus offering new insight into the Pauline letters.

Authorship of the Book of Hebrews

Addressed to Jewish believers, the book of Hebrews lays out

Table 5.
Clusters of verse segments of pauline letters (To save space, we used abbreviation. Say, Corinthians1_10_2-3 stands for First Corinthians verse segments 2 - 3 in chapter 10.).
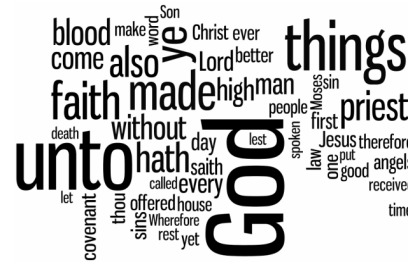
Cluster Number	Segments in One Cluster	Subjects
1	Corinthians1_8_1 Corinthians1_10_2-3	Food from idol feasts and the Lord's supper
2	Corinthians1_11_1	Propriety in worship
3	Corinthians1_11_2	Propriety in worship, Lord's supper
4	Corinthians1_11_3	Lord's supper
5	Corinthians1_12_2	One body, many parts
6	Corinthians1_12_3	One body, many parts
7	Corinthians1_13_1	Love
8	Corinthians1_14_1-3	Gifts of prophecy and tongues
9	Corinthians1_15_2	The resurrection of the dead
10	Corinthians1_15_4	The resurrection body
11	Corinthians1_15_5	The resurrection body
12	Corinthians1_1_2-3 Corinthians1_2_1	Wisdom of God and wisdom of the spirit
13	Corinthians1_3_2	On divisions in the church
14	Corinthians1_7_1-4	Marriage
15	Corinthians1_9_2	The right of an Apostle
16	Corinthians2_11_3	Paul boasts about his sufferings
17	Corinthians2_3_1-2	The glory of the new covenant
18	Galatians_4_3	Hagar and Sarah
19	Colossians_1_1 Corinthians1_1_1 Corinthians2_1_1 Ephesians_1_1 Galatians_1_1 Philemon_1_1 Philippians_1_1 Romans_1_1 Timothy2_1_1	Opening statement of each letter
20	Romans_11_2	Ingrafted branches
21	Romans_11_3	All Israel will be saved
22	Romans_14_1	The weak and the strong
23	All remaining segments of Paul's letters (158 segments)	Largest group containing various subjects
24	Corinthians1_16_2 Romans_16_1-2	Greetings to individuals
25	Romans_2_3	The Jews and the law
26	Galatians_3_1-2 Romans_4_1-2	Law, faith, promise of God, Abraham justified by faith
27	Romans_5_2	Death through Adam, life through Christ
28	Romans_6_1	Dead to sin, alive in Christ
29	Romans_6_2	Slaves to righteousness
30	Galatians_2_2 Romans_2_2 Romans_3_2 Romans_7_1-3 Romans_8_1	Law, Jews, righteousness, justification, condemnation
31	Thessalonians1_5_1	Coming of the Lord
32	Timothy1_3_1	Overseers and deacons
33	Titus_2_2	What must be taught to various groups

the superiority, perfection, and completion of Jesus Christ as mediator of God's grace for salvation. Jesus is the perfect high priest who offers Himself as a sinless sacrifice, and through Him we can receive eternal life. He is better than the angels and the forefathers of Israel. The letter enumerates the faith heroes in the history of Hebrews to demonstrate the overwhelming evidence of God's faithfulness, and therefore calls us to live by faith as the saints in the Old Testament did.
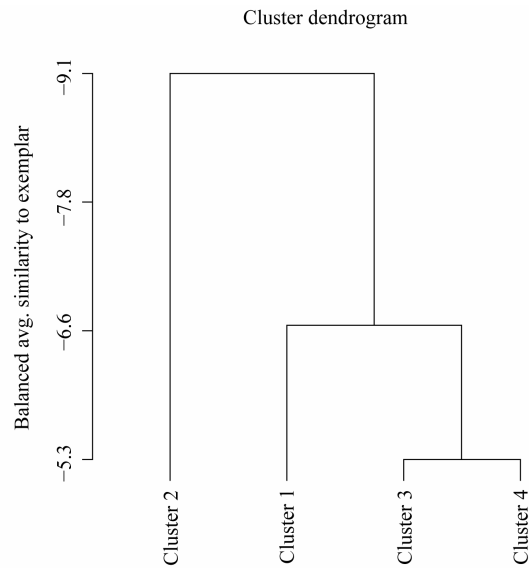
In all Pauline letters, there were a collection of his favorable words, "God", "Christ", "Lord", and "Jesus", as seen from the word clouds in **Figure 1**. But here in **Figure 3**, only the word "God" plus the words "faith" and "priest" are very visible. This discrepancy was certainly not a proof of authorship, but displaying some writing style difference between the 13 Pauline letters and the book of Hebrews.

To compare the book of Hebrews with the 13 Pauline letters, we clustered these 14 letters. The outcome showed that the letter to Hebrews formed a single letter cluster and the 13 Paul's letters appeared in 3 clusters (**Table 6**). The dendrogram in **Figure 4** also confirmed that Hebrews was quite different from the 13 letters of Paul in terms of their contents.

To study the authorship of the letter to Hebrews, we employed the 13 Pauline letters and 13 of non-Pauline books in the New Testament as training data to build our predictors. Using 10 fold cross validation, we selected 9 classifiers in Weka (http://www.cs.waikato.ac.nz/ml/weka/) that had less than 4 errors, indicating that they were highly accurate for this kind of prediction. The confusion matrices of these classifiers are presented in **Table 7** along with their predictions of the authorship of the book of Hebrews. Out of the 9 classifiers, only one predicted Paul as the author of the letter to Hebrews. By consensus we therefore could conclude that these 9 classifiers collectively predicted Hebrews as non-Pauline. It was worth noting that four of the classifiers, SMO, SL, IBk, and Bagging, tended to predict non-Pauline as Pauline as shown from their predictions of the training data in **Table 7**. But they all predicted Hebrews as

Figure 3.
Word clouds of top 50 frequent words in the letter to Hebrews.

Figure 4.
Dendrogram of the clusters shown in **Table 5**.

Table 6.
Clusters of Pauline letters and the book of Hebrews.

Cluster Number	Exemplar Letter Name	Letters in One Cluster
1	Colossians	Colossians, Corinthians1, Corinthians2, Ephesians, Galatians, Philemon, Philippians, Romans
2	Hebrews	Hebrews
3	Thessalonians2	Thessalonians1, Thessalonians2
4	Timothy1	Timothy1, Timothy2, Titus

Table 7.
Confusion matrices of 9 classifiers and their predictions of authorship of the letter to Hebrews.

Classified as	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b	a	b
Paul = a	13	0	11	2	13	0	13	0	13	0	12	1	13	0	12	1	13	0
Non-Paul = b	1	12	2	11	2	11	2	11	1	12	2	11	2	11	3	10	1	12
Classifier	SMO		SL		Bayes Net		NB		IBk		Bagging		RBF Network		RF		BLR	
Prediction of Hebrews Authorship	Non-Pauline		Non-Pauline		Non-Pauline		Non-Pauline		Pauline		Non-Pauline		Non-Pauline		Non-Pauline		Non-Pauline	

BLR = Bayesian Logistic Regression, SL = Simple Logistic, NB = Naïve Bayes, RC = Random Committee, RF = Random Forest, RT = Random Tree.

non-Pauline. The one classifier that predicted Hebrews as Pauline was a predictor that favored Paul as the author of other known non-Pauline letters (**Table 7**).

Conclusion

Paul is the most prolific New Testament writer that contributed 13 books to the collection of 27 books in the New Testament. He played a key role in the formation of early Christian faith and the church. In this work, we presented a computational approach to study of Paul's letters. Our goal was to employ the topic modeling and AP clustering algorithm to catalog the topics discussed either once or repeatedly in all his writings. The distributions of the topics of each document were calculated and used as features, to which the AP clustering algorithm was applied. This method provided a means to study Paul's messages as a whole at three granularities: letters, chapters, and verses, offering different views of the Pauline letters.

Our computational analysis implied all Pauline letters could be divided into six groups, well in line with the church tradition. They were First and Second Thessalonians, First and Second Corinthians, Colossians and Ephesians, Galatians and Romans, Philemon and Philippians, and First and Second Timothy and Titus. We were able to identify the topics covered either once or a few times in his letters. Among all letters, First Corinthians had the most unique chapters, whereas First Corinthians and Romans had the most unique verses or topics. Moreover, similar verses and chapters from different letters were also discovered. The grouping of the Pauline letters, their chapters, and their verses provided a global view of all his writings at different resolutions, with each complementing the other.

We also studied the authorship of the letter to Hebrews, an anonymous letter in the New Testament. Using known Pauline and non-Pauline letters as training data, we built nine highly reliable predictors, which collectively predicted that the book of Hebrews was not authored by Paul.

Acknowledgements

REFERENCES

Blei, D., Ng, A., & Jordan, M. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research, 3,* 993-1022.

Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science, 315,* 972-976. doi:10.1126/science.1136800

Goodspeed, E. J. (1933) *The meaning of Ephesians.* Chicago: University of Chicago Press.

Griffiths, T., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 5228-5235. doi:10.1073/pnas.0307752101

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. T. (2007). Topics in semantic representation. *Psychological Review, 114,* 211-244. doi:10.1037/0033-295X.114.2.211

Kostenberger, A. J., & Wilder, T. L. (2010) *Entrusted with the Gospel: Paul's theology in the pastoral epistles.* Nashville, TN: B&H Publishing Group.

O'Brien, P. T. (1999) The letter to the Ephesians (pillar new testament commentary). Grand Rapids, MI: Wm. B. Eerdmans Publishing Company.

Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Latent semantic analysis: A road to meaning.* Hillsdale, NJ: Laurence Erlbaum.

Stott, J. (2001) *The message of Romans: God's good news for the world.* Downers Grove, IL: IVP Academic.