Synoptic Network Analysis of the Four Gospels:
Why are there Four Gospels in the Bible?

Hajime Murai and Akifumi Tokosumi
Department of Value and Decision Science, Tokyo Institute of Technology
2-12-1, Ookayama, Meguro-ku, Tokyo 152-8552 Japan
h_murai@valdes.titech.ac.jp

Abstract — **There are many religions, such as Christianity, that have sought to spread their messages and have subsequently created a collection of documents. However, as the literature grows, it becomes more problematic to interpret any single text and perceive how it relates to other documents. This situation is common in other areas of human activity, not just religions.**

In this paper, we propose a method of Synoptic Network Analysis, to represent the relationships between multiple overlapping texts as a network, and to analyze the structure and semantics of the texts by clustering the network. This method is applied to the four traditional Christian Gospels in order to extract the main messages common to these Gospels and to highlight the dogmatic characteristics of each Gospel.

Unlike traditional literature-based approaches, Synoptic Network Analysis is a scientific method that incorporates falsifiability and replicability. The method makes it possible to scientifically evaluate various literature-based hypotheses.

I. Introduction

A. Problems for Various Canonical Texts

Recently, the Gospel according to Judas has attracted a great deal of attention [1]. The Judas Gospel shows that many Gospels may have been written to recount the acts and teachings of Jesus Christ. So, the traditional Gospels of Matthew, Mark Luke, and John are not the only Gospels ever written.

However, this revelation raises many questions about the Gospels, such as "how were they made and in what order?" "How are the Gospels related to each other?" "Why did the ancient church acknowledge four Gospels rather than selecting just one to convey their religious teaching?" and "why were other Gospels excluded from the traditional Bible?" These are questions that researchers have been tackling since ancient times.

There are undoubtedly many cases where a group of people have sought to spread their message and therefore developed a literature of "canonical documents," but where issues concerning the interpretations of the texts and the relationships between various individual documents have become problematic. This kind of situation exists not only within Christianity but within other religions and schools of political thinking. Such interpretative issues would seem to have direct influences on many matters in the modern world.

B. What is Synopsis?

The term 'synopsis' is derived from the Greek word Syn-Opsis (together - see) and originates in the charts those make it possible to "see together," for comparative purposes, the three very similar Gospels of Matthew, Mark, and Luke. These three Gospels are referred to as the Synoptic Gospels. While the synopsis format is believed to have been originally created in the fourth century, the first modern synopsis was produced in the eighteenth century.

In addition to a parallel chart of overlapping sections, the modern synopsis also represents correspondences between words and related sections. It is now regarded as an essential tool for the interpretation of "canonical texts" [2].

Recently, the term "synopsis" is being used in a less restrictive manner to refer to other comparisons beyond the three traditional synoptic Gospels. For example, many synopses also compare John with the three synoptic gospels, and some even add Thomas, which is not a traditional Gospel, in creating a chart of five parallel "canonical texts" [3].

C. Synopsis Research

There have been numerous studies concerning the synopsis and the synoptic Gospels throughout history, with a central focus of many relating to the order and the dependencies of the three synoptic Gospels, which are known as the Synoptic Problem. Traditionally the kind of problem, which is an issue for various "canonical texts," has been investigated using general literature–based approaches. However, a major shortcoming with such approaches is that they do not provide for hypotheses to be falsified or replicated, so it is very difficult to provide conclusive evidence to settle the debates one way or another.

More recently, a few studies are utilizing modern information technology in order to analyze the statistical characteristics of texts and to resolve the synoptic problem scientifically [4][5].

However, those studies have only been concerned with the chronological aspects to the construction of the Gospels. Thus, they have not addressed the synchronic semantic problems of how the central teachings of Christianity are dependent on four Gospels as a set of "canonical texts," rather than on just one.

II. The Aims of this Study

The central aim of this study is to develop a scientific information-technological method to analyze semantic differences that arise between multiple overlapping "canonical texts". We believe that this method can be applied not just to the Bible but also to the interpretation of the systematic thinking embodied within collections of "canonical texts" in other spheres.

Specifically, the goals of this study are to develop a method that can analyze:

- How do the central messages emerge through the existence of multiple overlapping "canonical texts"?
- How are these central messages modified under different combinations of the "canonical texts"?
- How does each text contribute to the construction of the central messages?

And, to apply this method to the four Synoptic Gospels included in the traditional Bible and to discuss the results obtained. Through these aims, this study illustrates numerically which messages Christianity has sought to convey with the selection of the four traditional Gospels.

III. Synoptic Network

A. Network Representation

Various methods have been utilized to represent human cognition throughout history. Recently, network representation seems to be one of the most promising methods.

There are many examples of network representations in various fields, such as semantic networks in linguistics and cognitive science [6], Social Network Analysis in sociology [7], the analysis of link structures that represent relations between WWW pages [8], and citation analysis used for analyzing scientific papers [9].

Network representation has the merit that network structures can represent complex relationships, making network representations particularly useful with handling the intrinsically complex nature of human concepts, such that exist in systematic thoughts. The effectiveness of network representation has been demonstrated in a study that focused on the Bible and Christian thoughts by analyzing a co-citation network constructed from documents by prominent theologians and extracting their characteristics [10].

B. Relationships between Parts of the Synoptic Texts

The internal structures of the Gospels are divided into segments that are called pericopes. Pericope is an ancient Greek words meaning cut-out. Each pericope corresponds to a small segment of a Biblical story that was transmitted orally.

In a synopsis, pericope units are identified for each Gospel and the corresponding pericopes in each Gospel are arranged in parallel. However, a particular pericope in one Gospel may correspond to multiple pericopes in another Gospel.

This one-to-many relationship is due to the editing process, as each Gospel writer combined pericopes that he believed to be related. Thus, if one author saw a connection between one pericope to several others, that particular pericope unit would be repeated in a number of sections within the Gospel. Accordingly, there are many pericopes in the four Gospels that have the same verses, because they were taken and edited from the same source pericope. In other words, the relationships between pericopes can be identified by matching common verses in the various Gospels.

For example, Fig. 1 shows the relationships for pericope No. 235 as depicted in a "Synopsis of the Four Gospels" [2].

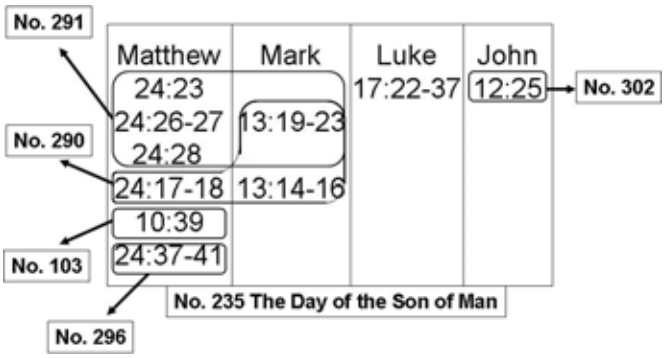

Fig. 1 Example of pericope relationships

As Fig. 1 shows, pericopes containing verses in common with pericope No. 235 are Nos. 103, 290, 291, 296 and 302. This suggests that the writer of Matthew perceived some relationships between pericopes 235, 103, 290, 291 and 296. Similarly, the writer of Mark imagined relationships to 290, 291, while the writer of John make a link between pericopes 235 and 302.

C. Constructing the Synoptic Network

The strategy used in constructing this network based on a single synopsis is to regard pericopes as nodes and the relationships between pericopes as edges (in fig. 1, the pericope 235 would be a node with edges to other nodes indicated by the arrows).

Different synopses will differ slightly in terms of

both of how pericopes are defined and their correspondences due to author interpretation. This study uses the "Synopsis of the Four Gospels" [2] based on Nestle Aland's "Greek New Testament" (version 26), as this is believed to be the basis for various synopses that have been used as data source in defining pericopes and their relationships. Tables 1 and 2 present numerical data relating to the synoptic network.

Nodes (Size)	367
Connected	273
Free	94
Edges	471

Table 1. Parameters of the synopsis network

Size of subgraph	Number of subgraphs	Total nodes
1(Free)	94	94
2	18	36
3	4	12
4	4	16
5	2	10
199	1	199

Table 2. Subgraphs in the synopsis network

VI. The Significance of the Four Gospels

A. Clustering and Extracting the Core

Although the maximum connected subgraph for the computed synoptic network contains the majority of nodes, this fact makes it difficult to identify its internal structure. In order to identify the internal structure, the maximum connected subgraph was clustered and the core element was extracted.

The procedure for clustering was as follows:

1. Detect cliques (sections where all nodes are completely interconnected) within the maximum connected subgraph (nodes connected with double lines in fig. 2).
2. Detect nodes of cliques that are shared by more than one clique.
3. Combine cliques sharing a node and form a cluster (colored area in fig. 2).

The feature of this algorithm is the ability to change cluster size by setting the size of cliques to be detected (i.e., more than three). If a created cluster is too large and its internal structure is not clear, then the algorithm can be reset to form a cluster by using larger cliques.

In this study, a detected cluster consisting of three or more cliques is regarded as a normal cluster. The maximum normal cluster is again divided according to cliques that are larger than four, which are called core clusters. This clustering algorithm is shown in Fig. 2.

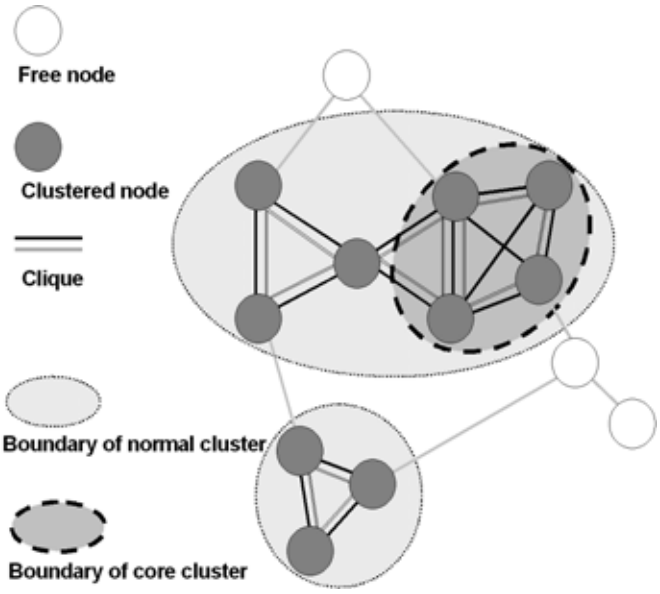

Fig. 2. Clustering method

The clustering algorithm is applied to connected subgraphs. In terms of node connection strengths, the strength of a subgraph is weaker than that of a normal cluster, which in turn is weaker than a core cluster. Thus, normal clusters are extracted from the maximum subgraph and core clusters are extracted from the maximum normal cluster.

The result of applying this clustering algorithm to the synoptic network for the four Gospels was that eight normal clusters were extracted. A second process of clustering applied to the maximum normal cluster (size 57) extracted four core clusters. Table 3 presents data concerning these clusters and their sizes.

	number	maximum
Subgraph	123	199
Normal clusters	8	57
Core clusters	4	8

Table 3 Subgraphs and clusters

Fig. 3 shows the clustered maximum connected subgraph of the synopsis network for the four traditional Gospels. This graph is made by Graphviz.

The elements of the core clusters are shown in Table 4, where the numbers indicate pericope reference number in Nestle-Aland [2].

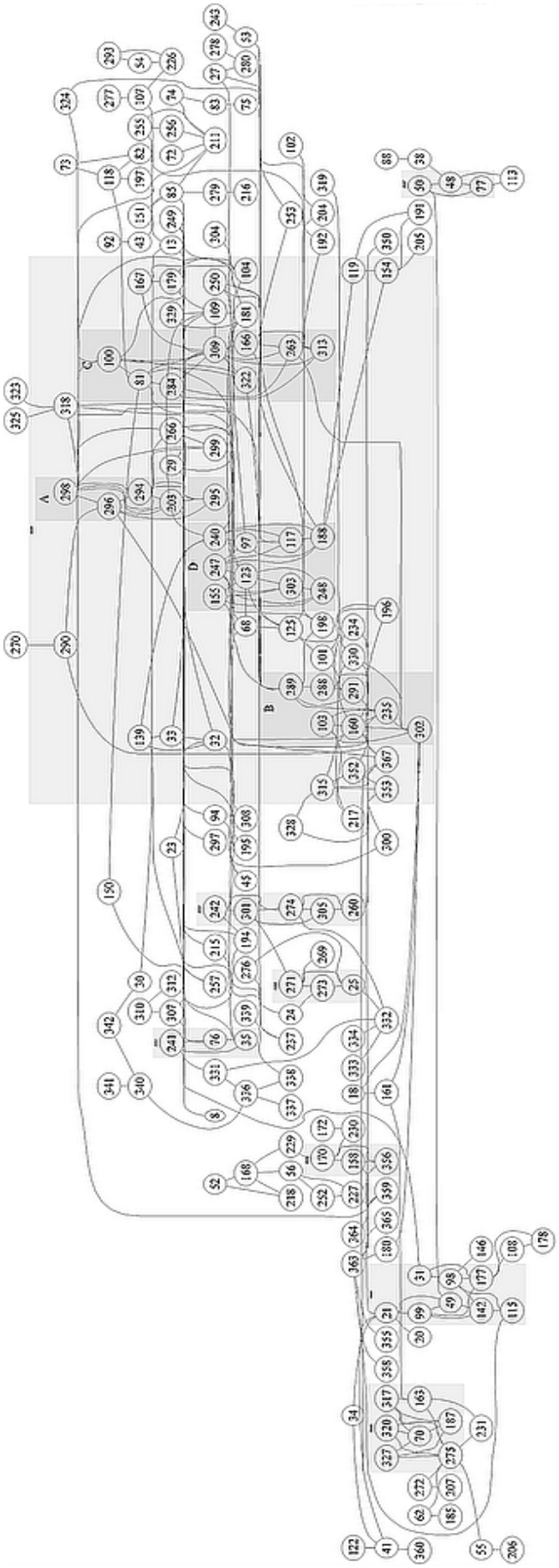

Fig. 3 Clustered maximum connected partial graph

Core Cluster	Pericope Number
A	203, 294, 295,296, 298
B	103, 160, 235, 288, 289, 291, 302
C	81, 100, 166, 263, 284, 309, 313, 322
D	97, 117, 188, 240, 247

Table 4 Elements of the 4 core clusters

B. Comparisons of Gospel Combinations

Four core clusters were extracted from the synoptic network based on the four Gospels. These core clusters occupy the central position of the network. This indicates that the core clusters represent the central concerns of the Gospel writers and the messages that they sought to convey.

In order to investigate how the central messages would differ according to the combination of the traditional Gospels, four further synoptic networks were created based on the different possible combinations of just three of the Gospels. The resultant core clusters and elements are shown in Tables 5 – 8.

A', B', and C' indicate the partial core clusters that lack some elements compared to the clusters A, B, and C.

Core Cluster	Pericope Number
A	203, 294, 295, 296, 298
B	103, 160, 235, 288, 289, 291, 302

Table 5 Core clusters of Matthew + Mark + Luke

Core cluster	Pericope Number
A'	203, 294, 296, 298
B'	103, 160, 235, 302
C	81, 100, 166, 263, 284, 309, 313, 322

Table 6 Core clusters of Matthew + Mark + John

Core Cluster	Pericope Number
A	203, 294, 295, 296, 298
B	103, 160, 235, 288, 289, 291, 302
C	81, 100, 166, 263, 284, 309, 313, 322
D	97, 117, 188, 240, 247

Table 7 Core clusters of Matthew + Luke + John

Core cluster	Pericope Number
B'	103, 160, 235, 302
C'	81, 100, 166, 263, 309, 313, 322
D	97, 117, 188, 240, 247

Table 8 Core clusters of Mark + Luke + John

While another set of synoptic networks was created by combining just two Gospels together, only the

combinations of Matthew + John and of Luke + John generate the core clusters of B' and C'. Synoptic networks based on only one Gospel failed to generate any core clusters.

These results indicate that the selection of four overlapping, but variant texts as a canon has been extremely effective in highlighting the core messages through a intricate network of interrelatedness created by the common verses.

Specific insights that can be gained from Tables 5-8 are as follows:

- Mark has the characteristics of being the greatest common divisor of the four Gospels. The fact that it does not contribute to the generation of a core cluster suggests that it has no unique central message.
- Cluster A does not emerge when Matthew is not included, suggesting that it is a characteristic of Matthew. However, pericope 295 seems to be dependent on Luke.
- As Cluster B is universal to all combinations it would appear to represent some message common to all four Gospels. However, pericopes 288, 289, 291 seem to be dependent on Luke.
- Cluster C does not emerge when John is not included, suggesting that it is a characteristic of John. However, pericope 284 seems to be dependent on Matthew.
- Cluster D is dependent on both Luke and John.

The relationships of the four Gospels and the core clusters are shown in Fig. 4.

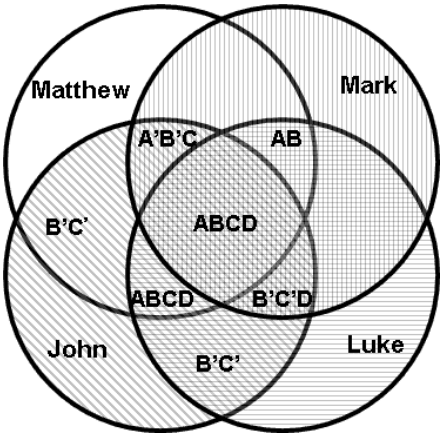

Fig. 4 Relations of the 4 Gospels and core clusters

C. What do the four clusters represent?

The theological significance of these clusters can be interpreted by looking at how the constituent pericopes match to sections of the Bible.

Cluster A includes the verses Mt25:1-13 and Mr13:33-37, which relate to the teaching that we should always be prepared for the Day of Judgment, because we do not know when it will come. Those verses are advocating an eschatological preparation of mind. Cluster A would seem to be dependent on Matthew, which has traditionally been seen as an attempt to evangelize none Christian Jewish people. Matthew emphasizes the urgency of becoming a Christian.

Cluster B is also concerned with eschatology (e.g., Lk17:22-37), but it also foretells of the persecution (Mr13:9-13) and recommends the path of discarding everything (Mt10:37-39). This cluster may be interpreted as providing concrete preparation guidance for the Day of Judgment. As a universal cluster, clearly all the Gospel writers were concerned about preparing for the Day of Judgment.

Cluster C appears to be related to teachings to the community of disciples, such as serving each other (Mt18:1-5) and not judging others (Lk6:37-42). It is a focus in John, which is believed to be the latest of the Gospels. It is thought that the focus on the preservation of the church and resolving inner problems emerged when the Gospel according to John was written.

Cluster D seems to be related to the Beelzebub Controversy (Lk11:14-23) over whether miracles were due to demons or not. While this cluster emerges in John and Luke, the Gospels of Matthew and Mark appear to be less interested in demons.

There has been much debate about what are the central messages of the Gospels. The results from the present statistical analyses of the computed synoptic networks indicate that preparing for the Day of Judgment (A, B) and promoting proper attitudes in disciples (B, C) were the central messages of the Gospel writers.

The analysis of multiple synoptic networks based on different combinations of the four Gospels indicates that the same core clusters still emerge when the Gospel according to Mark is excluded. This finding is consistent with the results of previous Biblical studies that suggest that Mark is probably the basis for all four Gospels.

On the other hand, in the traditional canonical interpretation such as "Catechism of The Catholic Church" [11], eschatological teaching is not thought as important messages of Jesus Christ in Catholic Church. Famous historical theologians' common central dogma is salvation through incarnation of the God [10] rather than eschatological teaching.

Traditional literary-based approaches have generated many hypotheses about the central messages of the four Gospels. However, because these literary-based approaches do not provide methods of objectively evaluating and testing their hypotheses, this study proposes a synoptic network analysis method that makes it possible to scientifically and objectively detect the central messages of the synoptic Gospels.

V. Conclusion

- The method of synoptic network analysis has been proposed that creates a network from multiple overlapping texts and extracts their central messages.
- Synoptic network analysis differs radically from traditional literary-based approaches, by incorporating methods of falsification and replicability, which are essential in order to scientifically evaluated and test various hypotheses about the central messages of a canon.
- Synoptic network analysis was applied to the traditional four Gospels and four central messages were detected. This study also investigated how the central messages change according to the combination of Gospels, as well as examining the interdependencies between the various messages and the different Gospels.

This study has employed the unit of the pericope which is used in general Biblical studies. However, the method developed in this study utilizes the pericope units to yield more scientific and objective results.

The developed synoptic network analysis method can also be applied to other textual canons, apart from the Bible, that consist of multiple, overlapping texts that form a single body of systematic thought. The method is particularly relevant for religions, such as Buddhism and Islam, and for schools of political and social ideology that can draw on a rich tapestry of discourse, as well as fables and myths that have been passed down through various manuscripts.

[1] Rodolphe Kasser, The Gospel of Judas, National Geographic, 2006.
[2] Kurt Aland, Synopsis of the four Gospels revised standard version, the United Bible Societies, 1982.
[3] Migaku Satoh, The Japanese Synopsis of the Gospels in Colors, Iwanami, 2005.
[4] Maki Miyake, Hiroyuki Akama, Masanori Nakagawa, Nobuyasu Makoshi, The Computed Synoptic Table--Tele-Synopsis for Biblical Research, Proceedings of ACH/ALLC 2005, pp.152~154
[5] A statistical approach to the synoptic problem, http://www.davegentile.com/synoptics/main.html
[6] George A. Miller, WordNet: a lexical database for English, Communications of the ACM, vol. 38, pp. 39-41, 1995
[7] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks, PHYSICAL REVIEW E, vol. 69, 2004
[8] R. Lempel and S. Moran, SALSA The Stochastic Approach for Link Structure Analysis, ACM Transactions on Information systems, pp. 131-160, 2001.
[9] Liu Mengxjong, Progress In Documentation The Complexities of Citation Practice: A Review of Citation Studies, Journal of Documentation, vol. 49, pp. 370-408, 1993.
[10] Hajime Murai and Akifumi Tokosumi, A Network Analysis of Hermeneutic Documents Based on Bible Citations, XXVII Annual Conference of the Cognitive Science Society, pp 1565-1570, 2005.
[11] John Paul II, Catechism of the Catholic Church, http://www.vatican.va/archive/ccc/index.htm, 1992.

Appendix
Correspondence between pericope reference numbers and Bible verses
(Only main elements, referential verses are omitted).

No.	Core cluster A
203	Lk12:35-48
294	Mr13:33-37
295	Lk21:34-36
296	Mt24:37-44
298	Mt25:1-13
No.	**Core cluster B**
103	Mt10:37-39
160	Mt16:24-28/Mr8:34-/Mr9:1/Lk9:23-27
235	Lk17:22-37
288	Mt24:3-8/Mr13:3-8/Lk21:7-11
289	Mt24:9-14/Mr13:9-13/Lk21:12-19
291	Mt24:23-28/Mr13:21-23
302	Jn12:20-36
No.	**Core cluster C**
81	Lk6:37-42
100	Mt10:17-25
166	Mt18:1-5/Mr9:33-37/Lk9:46-48
263	Mt20:20-28/Mr10:35-45
284	Mt23:1-36/Mr12:37-40/Lk20:45-47
309	Jn13:1-20
313	Lk22:24-30
322	Jn15:18-25
No.	**Core cluster D**
97	Mt9:32-34
117	Mt12:22-30/Mr3:22-27
188	Lk11:14-23
240	Jn7:14-39
247	Jn8:48-59