Cahit Arf: Exploring His Scientific Influence Using Social Network Analysis and Author Co-citation Maps

Yaşar Tonta[1] and A. Esra Özkan Çelik[2]

[1] *tonta@hacettepe.edu.tr*
Department of Information Management, Hacettepe University, 06800 Beytepe, Ankara, TR

[2] *esra@hacettepe.edu.tr*
Registrar's Office, Hacettepe University, 06800 Beytepe, Ankara, TR

Abstract

Cahit Arf (1910-1997), a famous Turkish scientist whose picture is depicted in one of the Turkish banknotes, is a well known figure in mathematics with his discoveries named after him (e.g., Arf invariant, Arf rings, the Hasse-Arf theorem). Although Arf may not be considered as a prolific scientist in terms of number of papers (he authored a total of 23 papers), his influence on mathematics and related disciplines was profound. As he was active before, during and after the World War II, Arf's contributions were not properly listed in citation indexes and thus did not generate that many citations even though several papers with "Arf" in their titles appeared in the literature. This paper traces the influence of Arf in scientific world using citation analysis techniques first. It reviews the scientific impact of Arf by analyzing both the papers authored by Arf and papers whose titles or keywords contain various combinations of "Arf rings", "Arf invariant", and so on. The paper then goes on to study Arf's contributions using social network analysis and author co-citation analysis techniques. CiteSpace and pennant diagrams are used to explore the scientific impact of Arf by mapping his cited references derived from Thomson Reuters' Web of Science (WoS) database. The paper ends with a discussion of whether data analysis techniques used in this study can be useful to study the scientific impact of researchers retrospectively.

Key words: Cahit Arf, author co-citation analysis, social network analysis, pennant diagrams, CiteSpace.

Introduction

Cahit Arf is a world-renowned Turkish mathematician who had had significant inventions to his credit, which still are in use today such as "Arf invariant", "Arf rings" and "Arf closure". He was born in Selanik (Thessaloniki) on February 18, 1910. With the outbreak of the Balkan Wars, his family migrated to Istanbul in 1912, then to Ankara, and finally settled in Izmir. As a brilliant student, Arf had successfully completed the École Normale Supérieure in Paris in two years and then worked as a teacher at Galatasaray High School for a year with great willingness. In 1933, he joined the Mathematics Department of Istanbul University as an assistant professor and finally began to work as a mathematician for academic purposes. In 1937, Arf went to Göttingen to have his PhD degree under the supervision of Professor Helmut Hasse. He completed his doctoral studies in one and a half years (1938); the Hasse-Arf theorem was an outcome of his doctoral thesis. After completing his PhD, he stayed in Göttingen one more year at the request of Professor Hasse (O'Connor & Robertson, 1998) and began to study on the quadratic forms over a field of characteristic two to improve the theory established by Ernst Witt (1937). In 1941, Arf published the results of his study and completed the theoretical gap by introducing an important invariant of quadratic forms over a field of characteristic two (Arf, 1941). The invariant is well known as Arf invariant and is the key for the solution of several classical and fundamental problems about the topology of manifolds (Ikeda, 1998; Önder, 1990). In 1948, Arf published another significant contribution in the *Proceedings of the London Mathematical Society* (Arf, 1948). The most significant follow up to Arf's paper came from Lipman (1971) who was the first mathematician to mention "Arf rings" in the literature (Sertöz, 1997).

Besides his many significant inventions in mathematics, Professor Arf worked in Istanbul University until his involvement, upon invitation, in the foundation of the Scientific and Technological Research Council of Turkey (TUBITAK). He served as the head of TUBITAK's Science Council between 1963 and 1967. In 1963, he joined the Mathematics Department of Robert College in Istanbul and worked at the Institute for Advanced Studies in Princeton, New Jersey, for two years. While at Princeton, he was invited to spend a year at the University of California at Berkeley as a visiting scholar. Then he made his final return to Turkey, joined the Mathematics Department of the Middle East Technical University and continued his studies there until his retirement in 1980. He received the prestigious Inonu Award in 1948, TUBITAK Science Award in 1974, and Commandeur des Palmes Académiques (France) in 1994. Arf was a member of the Mainz Academy and the Turkish Academy of Sciences. He served as the president of the Turkish Mathematical Society between 1985 and 1989. After his many contributions to the international mathematics society and the Turkish scientific environment, he passed away on December 26, 1997 in Bebek, Istanbul, at the age of 87 (O'Connor & Robertson, 1998).[1] In 2009, the banknote of 10 Turkish Lira was issued with Professor Cahit Arf's portrait depicted on it.[2]

Within his 87 years life time, Arf was well known for the Arf invariant of quadratic forms over a field of characteristic two, Arf rings, Arf semigroups and the Hasse–Arf theorem, among others. This paper aims to study the influence of Cahit Arf's papers retrospectively by means of author co-citation analysis and social network analysis. Papers that referred to Arf's contributions in their titles and topics were identified and an author co-citation analysis was carried out. CiteSpace was used to find out Arf's place in mathematics and his impact on the basis of bibliometric analysis, author co-citation analysis and pennant diagrams (White, 2007a, 2009). We discuss the consequences of such an approach and conclude that author co-citation analysis and pennant diagrams can shed further light on the influence of authors and help put their works in full perspective.

Literature Review

Social network analysis (SNA) is used to study and visualize the structures of social networks. Based on graph theory, SNA has been widely used to reveal the relationships among documents, journals and authors (Otte & Rousseau, 2002). Scientific and intellectual ties and collaboration between researchers can be identified using bibliometric data (e.g., citations). The structure of scientific disciplines as social networks can be mapped by means of visualization software such as CiteSpace (Chen, 2006).

Author co-citation analysis (ACA) was first introduced by White and Griffith (1981). ACA assumes that the two researchers being cited together in the scientific literature are likely to share the same research interests and work in the same field. Researchers working in the same domain get clustered through ACA, thereby facilitating the discovery of social structure among researchers as well as among research domains. The outcome of ACA studies were used to map and visualize the structure of several scientific disciplines including information science and macroeconomics (see, for example, McCain, 1984, 1986, 1990; McCain, Verner, Hislop, Evanco, & Cole, 2005; White & McCain, 1998).

White recently combined author co-citation analysis with information retrieval (IR) and relevance theory (RT) to study the influence of a seed work, author or paper (White, 2007a,

[1] See also: http://en.wikipedia.org/wiki/Cahit_Arf.
[2] http://www.tcmb.gov.tr/yeni/banknote/E9/10tle.htm

2007b, 2009). He used the weights of term frequency (tf) and inverse document frequency (idf) formula to draw pennant diagrams for works co-cited with a seed work, the authors co-cited with a seed author or articles and books co-cited with a seed article. Originally, tf values are used in IR to determine the relevance of a term within a given document (the more a term is used in a single document, the higher its relevance to a given query) while idf values are used to determine its relevance within the entire document collection (the more a term is used in different documents in a collection, the less discriminatory -and therefore topically less relevant- it becomes). Sparck Jones (1972) created idf as a measure for weighting the "statistical specificity" of terms. The idf measure pushes the related term's weight down in the rankings so that terms that occur relatively frequently in the document collection are considered less relevant to a query.

Bibliometric data and IR techniques are used in pennant diagrams "to mimic a relevance theoretic model of cognition on the user side" (White, 2007a, p. 537). Pennant diagrams are scatterplots of tf values representing "cognitive effects" of works/authors/articles in the context of a seed work/author/article, and idf values representing the "processing effort" of the user. Cognitive effects (tf) and ease of processing (idf) of works, authors or articles determine their relevance to a seed work, seed author or seed article (White, 2007a, p. 550). Whereas tf*idf formula multiplies the two to come up with a single score of relevance, pennant diagrams plot tf values on the *x* axis and idf values on the *y* axis without multiplication (White, 2007a, p. 541).[3]

White used pennant diagrams innovatively to study the influence of a work (*Moby Dick*), an author (Howard D. White) and a paper (Stephen Harter's "Psychological Relevance and Information Science") and interpreted his findings based on ACA, IR and RT (White, 2007a). Pennant diagrams proved to be useful in discovering new relationships between works, authors and papers. Using both SNA and White's approach, the current study is an attempt to explore the scientific legacy of a world-famous Turkish mathematician, Cahit Arf, whose impact may not be measured readily by using citation indexes.

Method

We think the scientific legacy of Cahit Arf is underrepresented in Thomson Reuters' (formerly ISI's) citation indexes. Hence, we decided to address several research questions to paint a better picture of his accomplishments as a mathematician. For example, who were the authors being co-cited most often with Arf? How high were their h index scores? Can we trace the scientific influence of Arf through paper titles and topics that contain the terms "Arf invariant", "Arf rings" and so on? Which paper of Arf received the highest number of citations? Will the pennant diagram of Arf's most frequently cited paper provide further insight into his influence in mathematics as well as in other disciplines?

To address these research questions, we first searched for bibliographic records for Cahit Arf ("Arf C*") in the Web of Science (WoS). Search results were in no way satisfactory (for the reasons explained before) to study the influence of Arf's contributions. We then searched for records having "Arf*" in their titles and/or topics and analyzed them using CiteSpace, which is a freely available application developed by Dr. Chen to analyze the scientific literature and visualize the trends and patterns in the data.[4] We studied the distribution of citations on the

[3] See White (2007a) for a more comprehensive discsussion of theoretical foundation of bibliometrics, IR and RT along with their use in three different types of pennant diagrams for a seed work, seed author and seed article.
[4] http://cluster.cis.drexel.edu/~cchen/citespace/

co-citation network derived from CiteSpace by time-slicing them in 10-year intervals to identify other influential mathematicians who were co-cited with Arf.

We searched for cited references of Arf from WoS and found that his paper on what is later called "Arf invariant" were cited a total of 100 times (Arf, 1941). Then, using Arf (1941) as the seed work, we mapped author co-citation analysis results onto a pennant diagram using White's approach (White, 2007a, 2009). To create Arf's pennant diagram, we used the most highly cited 20% of references contained in 100 papers citing Arf (1941). Using Arf's pennant diagram, we identified other influential authors in the author co-citation network whose work was most relevant to those of Arf and, more specifically, to Arf's seed work (White, 2007a).

Findings and Discussion

Arf published a total of 23 papers between 1939 and 1966 (Arf, 1990). The distribution of these papers by language is as follows: 12 in French, 6 in German, 4 in English and 1 in Italian. The two of his English papers were listed in Thomson Reuters' citation indexes along with two English abstracts (Table 1). They were cited a total of four times, thereby making Arf's h index score 1. This is by no means commensurate with his fame, however.

Table 1. Arf's papers and abstracts listed in Thomson Reuters' citation indexes

1.	Arf, C., Imre, K. & Ozizmir, E. (1965). On algebraic structure of cluster expansion in statistical mechanics. *Journal of Mathematical Physics*, 6(8): 1179-&. (Times cited: 3)
2.	Arf, C. (1952). On methods of Rayleigh-Ritz-Weinstein. *Proceedings of the American Mathematical Society*, 3(2): 223-232. (Times cited: 1)
3.	Arf, C. (1951). On Rayleigh-Ritz-Weinstein method. *Bulletin of the American Mathematical Society*, 57(4): 269-270. (Times cited: 0)
4.	Arf, C. (1951). On a free boundary problem in elasticity. *Bulletin of the American Mathematical Society*, 57(2): 136-136. (Times cited: 0)

Note: The query "Arf C*" was run on Thomson Reuters' databases SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH using the all years time span (on July 25, 2011).

In spite of his many prominent papers with significant inventions named after him, Arf's bibliometric profile is almost nonexistent. None of Arf's now-classic papers that influenced advanced mathematics profoundly was indexed in Thomson Reuters' citation indexes. Uncitedness among mathematicians with Fields Medal (considered to be the Nobel Prize of mathematics) or even Nobel Laureates is not uncommon (Egghe, Guns & Rousseau, 2011). Yet, this is not the case for Arf. Even though his works were not listed in citation indexes properly, a cited reference search under "Arf C*" produced a total of 146 citations. Compare this with a total of four citations that Thomson Reuters used to calculate Arf's h index. Arf's one paper (1941) alone received a total of 100 citations (which is not listed in ISI indexes).[5] Similarly, Lipman wrote a paper entitled "Stable ideals and Arf rings" in 1971, which were cited a total of 94 times (Lipman, 1971). Lipman was referring to Arf's original paper that appeared in the *Proceedings of the London Mathematical Society* in 1948, which is not listed in ISI indexes, either (Arf, 1948). None of these direct or indirect citations to Arf's papers (1941, 1948) contributed to Arf's h index score. This is mainly due to the fact that Arf's

[5] American Mathematical Society's Mathematical Reviews (MR) Citation Database on the Web (http://www.ams.org/mathscinet/) provides 31 citations to Arf's six papers including 20 citations to Arf (1941). The MR database covers relatively current citations (year 2000 to present) (http://www.ams.org/mathscinet/help/citation_database_understanding.html).

significant contributions were published before the then ISI's citation indexes came into being and a great majority of them were not written in English journals and therefore not indexed by ISI. One could only speculate as to how many citations Arf's original papers would have generated had they been listed in ISI indexes.

Although Arf's papers were not properly listed in ISI's citation indexes and citations to them therefore did not count towards Arf's h index score, Arf's influence can be observed further through paper titles that contain various references to Arf's works (e.g., "Arf invariant", "Arf rings", "Arf-Hasse theorem" and so on). We performed an advanced search in Thomson Reuters' citation indexes[6] and found a total of 43 papers (38 articles, 4 proceedings paper, and 1 correction) with "Arf*" in their titles.[7] Note that not all 43 papers contained references to Arf's works. In fact, only 15 of them did (a total of 16 citations). Arf's classic papers (1941, 1948) received 8 and 5 citations, respectively. These 43 papers were mainly classified under Mathematics and were cited a total of 279 times (h index 9, max. citation per item: 95 - Lipman, 1971-, avg. citation per item: 6.49).

We performed a topical search (TS) in Thomson Reuters' citation indexes and found an additional 52 papers with "Arf*" in their keywords (i.e., topics).[8] Note that only 9 out of 52 papers cited Arf's two papers (1941, 1948). These 52 papers received a sum of 208 citations (h index 8, max. citation per item: 25, avg. citation per item: 4.00).

Altogether, 95 papers with "Arf*" in their titles or topics (e.g., abstract keywords or keywords given by their authors) published between 1965 and 2011 were cited a total of 487 times (h index 11, avg. max. citation per item: 95, avg. citation per item: 5.13). It should be noted that 24 authors cited Arf's works 25 times but the great majority of authors (69 out of 95) who referred to Arf's works in titles or keywords of their papers did not necessarily give due credit to Arf in their reference lists by properly citing Arf's papers. Apparently, they were either unaware of the existence of Arf's papers or they did not cite them because Arf's papers became a part of "regular scientific discourse" (Tonta & Darvish, 2010, p. 169). Indeed, as Terzioğlu (1998) points out, Arf's name is so intertwined with mathematics one needs to work hard to find citations to Arf's papers. Some authors using Arf invariant in their works seem to have referred to it as a mathematical symbol or notation (e.g., "Arf(X)") without, perhaps, thinking that these three characters are actually the last name of a Turkish mathematician (Tosun Terzioğlu, personal communication, August 16, 2011).
We know that ISI indexers tended to make "implicit" citations to works of art or musical scores within the arts and humanities papers "explicit" by indexing them accordingly (Al, Şahiner & Tonta, 2006, p. 1012; Garfield, 1980; Stern, 1983). Such citations count towards one's cited references and possibly towards his/her h index score. Yet, we are not aware of any Thomson Reuters convention that makes implicit references in paper titles or topics explicit, thereby giving credit to those whose works get cited tacitly. An implicit reference

[6] We used the following query: TI=(arf theorem) OR TI=(arf invariant*) OR TI=(arf ring*) OR TI=(arf propert*) OR TI=(arf filtration*) OR TI=(arf semigroup*) OR TI=(arf singularit*) OR TI=(arf equivalence) OR TI=(arf closure*). Irrelevant items were discarded.

[7] The terms used in the titles of these papers are as follows: Arf invariant* (the most common one) or Arf's invariant, Hasse-Arf filtrations, Arf rings, Arf numerical semigroups, Arf semigroups, Arf functions, Arf characteristics of singularities, Arf closure, Arf equivalence, Arf-Kervaire invariant, Hasse-Arf theorem, and Hasse-Arf property.

[8] Note that TS gets the keywords from titles, abstracts, and the author-assigned keywords. Therefore, we excluded the titles to find the unique items that would be retrieved only through keywords that come from abstracts or author-assigned keywords. In other words, these 52 papers do not overlap with the previous 43 papers that we identified through title search.

made explicit by indexers within a paper is equal to one citation. But what about an implicit reference found in the title or topic of a paper? It would certainly be worth more than one citation. For instance, Sertöz (1990) in his article on Arf rings cites no fewer than eight papers that were clearly based, at least in part, on Arf's work and the two of those papers have "Arf*" in their titles (excluding Lipman's oft-cited paper that we mentioned earlier).

Suppose that we decided to find out what would Arf's total influence be if we weighted the papers with "Arf*" in their titles (43) and keywords (52). Further suppose that each paper with "Arf*" in its title and keywords list is worth 10 and 5 citations, respectively. This would give us a total of 690 citations to Arf's works. One would safely claim that Arf's h index score would have been much higher if his works were listed in Thomson Reuters indexes and citations to his works in paper titles and keywords were made explicit.[9]

Author Co-citation Analysis

We further analyzed the papers with "Arf*" in their titles and topics (i.e., keywords) using CiteSpace. We selected the most frequently cited 20% of those papers and used author co-citation analysis to visualize them in 10-year time slices starting from 1966 (Fig. 1).

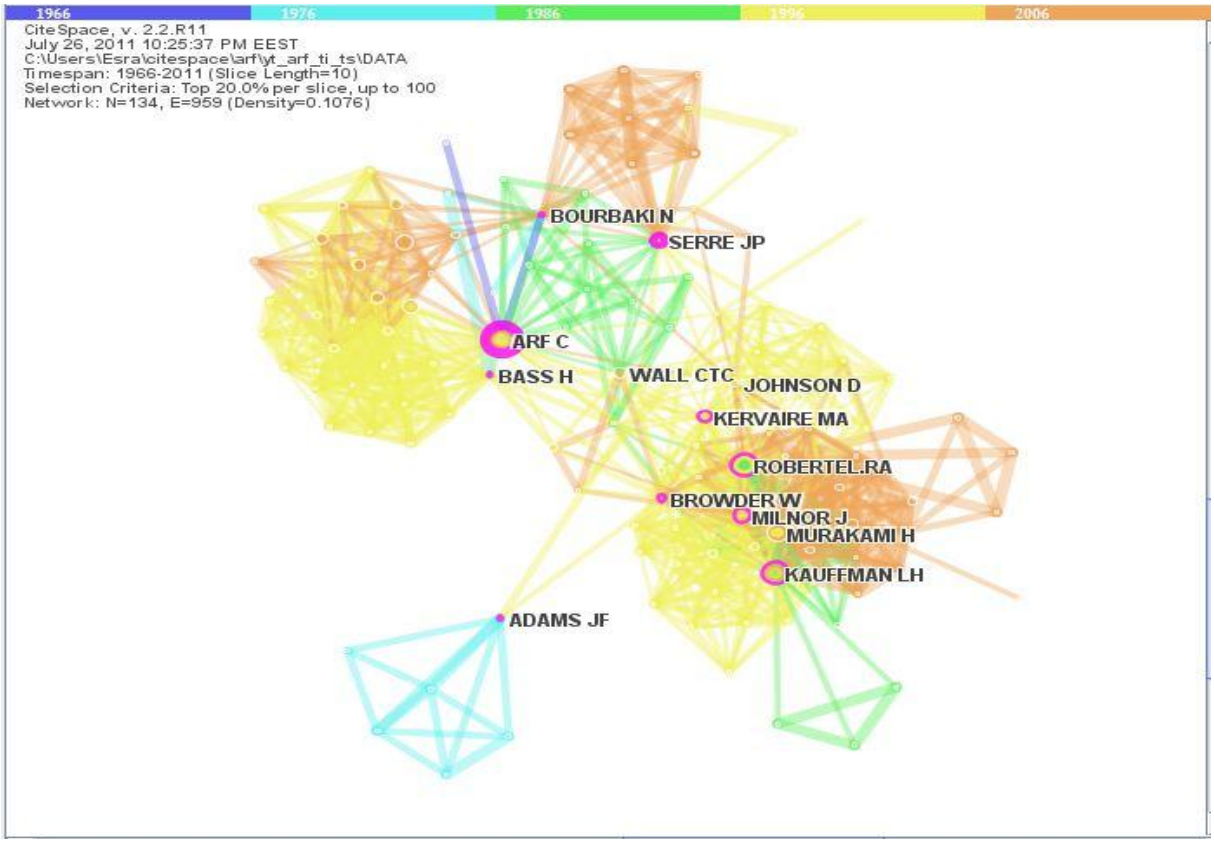

Fig. 1. Author co-citation network (co-cited with authors of papers with "Arf*" in their titles/keywords)

Clearly, Arf C is the centroid of such an author co-citation network. Kaufman LH, Robertello RA, Milnor J, Serre JP, Kervaire MA, Browder W, Bass H, Adams JF, and Bourbaki N are

[9] For comparison, J. Lipman wrote 34 papers (including the one on Arf rings receiving 94 citations) and were cited 622 times (h index 12).

the pivotal nodes in the network as the purple rings indicate. Pivotal nodes are strategically important in pulling together other nodes and therefore their betweenness centrality scores[10] are higher (Chen, Song, Yuan & Zhang, 2008). These pivotal nodes and many light colored ones not labeled represent very important mathematicians.

Orange colored clusters of the network indicate works that are published most recently (from 2006 to 2011). They make up a considerable part of the full network in proportion. If we add yellow colored parts of the network (representing the time span 1996 to 2005) to orange colored ones, we could obtain almost the whole network. In other words, Cahit Arf is still cited heavily in papers with "Arf*" in their titles or keywords, along with the great mathematicians of today.

As we indicated earlier, a cited reference search for "Arf C*" generated 146 results. Whereas his official h index is 1, his h index based on citations to his works that are not listed in citation indexes would be 5. Table 2 provides the names of 10 famous mathematicians who were co-cited frequently with Arf along with their h index scores.

Table 2. Co-cited author distributions

Co-cited Authors	*Frequencies*	*H Index Scores*
ARF C	146	1
SERRE JP	30	23
WITT E	23	(No ISI records under his name)
BOURBAKI N	21	(No ISI records under his name)
MILNOR J	20	29
DIEUDONNE J	20	11
SAH CH	20	14
KERVAIRE MA	18	9 (with respect to 9 paper)
OMEARA OT	18	11
WALL CTC	18	25
KNESER M	14	8

Names in Table 2 come from the most frequently cited 20% of Arf's co-cited authors.[11] For example, an award-winning mathematician Jean-Pierre Serre is on the top of the co-cited authors list with an h index of 23. He won Abel Prize (2003), Fields Medal (1954), Wolf Prize in Mathematics (2000) and Balzan Prize (1985).[12] He also gave a speech about Cahit Arf in "Cahit Arf Seminars" in 2006 at the Middle East Technical University in Ankara, Turkey. Milnor J has an h index of 29 and appeared in Arf's co-cited authors list, too. Witt and Kervaire are also on the above list. Witt, too, worked in Göttingen and is still being cited with his works on Witt algebra, Witt decomposition, Witt design (Witt geometry), Witt group, Witt index, Witt polynomial, Witt ring, Witt scheme, Witt's theorem, Witt vector, Bourbaki–Witt theorem, and Shirshov–Witt theorem.[13] However, he has no record with his name in citation indexes, either. "Arf invariant" is sometimes referred to as Arf-Kervaire invariant due to Kervaire's later work. A group of 20th century mathematicians published nine important

[10] Betweenness centrality score measures how nodes facilitate the flow in the network (Otte & Rousseau, 2002, pp. 442–443).
[11] It should be noted that "Nicolas Bourbaki" (Bourbaki N) is actually a pseudonym referring to more than one authors. For more information, see the Wikipedia article at http://en.wikipedia.org/wiki/Nicolas_Bourbaki.
[12] Wikipedia: http://en.wikipedia.org/wiki/Jean-Pierre_Serre
[13] Wikipedia: http://en.wikipedia.org/wiki/Ernst_Witt

books under the pseudonym of Nicolas Bourbaki,[14] presenting an exposition of modern advanced mathematics and receiving hundreds of citations to them. (Serre JP, Dieudonne J and Omeara OT also wrote influential textbooks.) It should be pointed out that names that appeared in author co-citation network and the top co-cited author distributions such as Serre, Milnor, Dieudonne, Kervaire, Kneser, Bass, Adams and Witt are all well known mathematicians.

We used co-citation statistics to draw a pennant diagram of Cahit Arf to trace his scientific influence further. What follows is a discussion of the influence of Cahit Arf based on the pennant diagram of his most significant work, "Untersuchungen über quadratische Formen in Körpern der Charakteristik 2" (Arf, 1941).

The outcome of the cited reference search under "Arf C*" (as of July 25, 2011) showed that his paper in which he first described the "Arf invariant" (Arf, 1941) were cited 100 times, constituting more than two thirds of all citations to his works. We used Arf (1941) as the seed paper and identified the most highly cited 20% of references contained in those 100 papers citing Arf (1941). We found that 234 papers were co-cited with Arf's seminal work (1941) at least once. To refine the resultant pennant diagram (otherwise it would be difficult to read the labels of nodes), we used 34 (out of 234) papers which were co-cited with Arf (1941) at least four times. Searching citation indexes, we found the tf (Items Ranked) and df (Items in File) values and used the following tf*idf formula (Manning & Schütze, 1999)

$$\text{weight}(i,j) = (1 + \log(tf_{i,j}))\log(N/df_i) \qquad (1)$$

where all term counts are greater than or equal to 1, logarithms are based 10, and *N* is the total number of items in the Thomson Reuters' data collection[15] (White, 2007a). Values for *x* and *y* axes in the pennant diagram in (Fig. 2) were calculated according to (1). (Appendix 1 gives the tf and idf values sorted by dfs and tfs.)

Arf with his most significant paper (1941) appears at the tip of the pennant (Fig. 2) at the right hand side as the seed term (White, 2007a, p. 541). (The years next to authors' names indicate publication years of relevant papers.) The seed term generates a bibliometric distribution which predicts the relevance of any associated term with itself. Higher tf and idf scores of any associated term will produce greater predicted relevance to the seed.

While the pennant diagram narrows through the right hand side, papers represented by points in Figure 2 become increasingly more relevant to that of Arf (1941). Authors in the left-most column were co-cited at least four times with Arf's paper (1941) while the ones to its right were co-cited progressively more. For instance, Witt's 1937 paper were co-cited 12 times with that of Arf (1941). (Figures in the *x* and *y* axes are both logged.) This is not a coincidence as Arf developed and completed Witt's work, as we explained earlier. Similarly, O'Meara's 1963 textbook *Introduction to Quadratic Forms* is the second highest co-cited work with Arf (1941) because it builds, at least in part, upon Arf's seminal work on the subject.[16] Thus, it is relatively easier to discern the increasing relevance of the papers as we move to the tip of the pennant.

[14] Wikipedia: http://en.wikipedia.org/wiki/Nicolas_Bourbaki
[15] The number of items in ISI file is assumed to be 5 million.
[16] Note that references to authors and their works in the pennant diagram can be found in Appendix 1.

White divided the pennant diagram into three (sections A, B, and C) and interpreted the results from various different angles such as the specificity of works (i.e., topicality), ease of processing, age and authority, fame and so on (White, 2007a). For instance, co-cited authors at the top (section A) of the pennant are topically more specific and therefore more relevant to that of Arf (1941) compared to the ones at the middle (section B) or bottom (section C) of the pennant. Similarly, co-cited authors in section A are juniors of Arf while the ones in Section B and C are his peers and seniors, respectively.

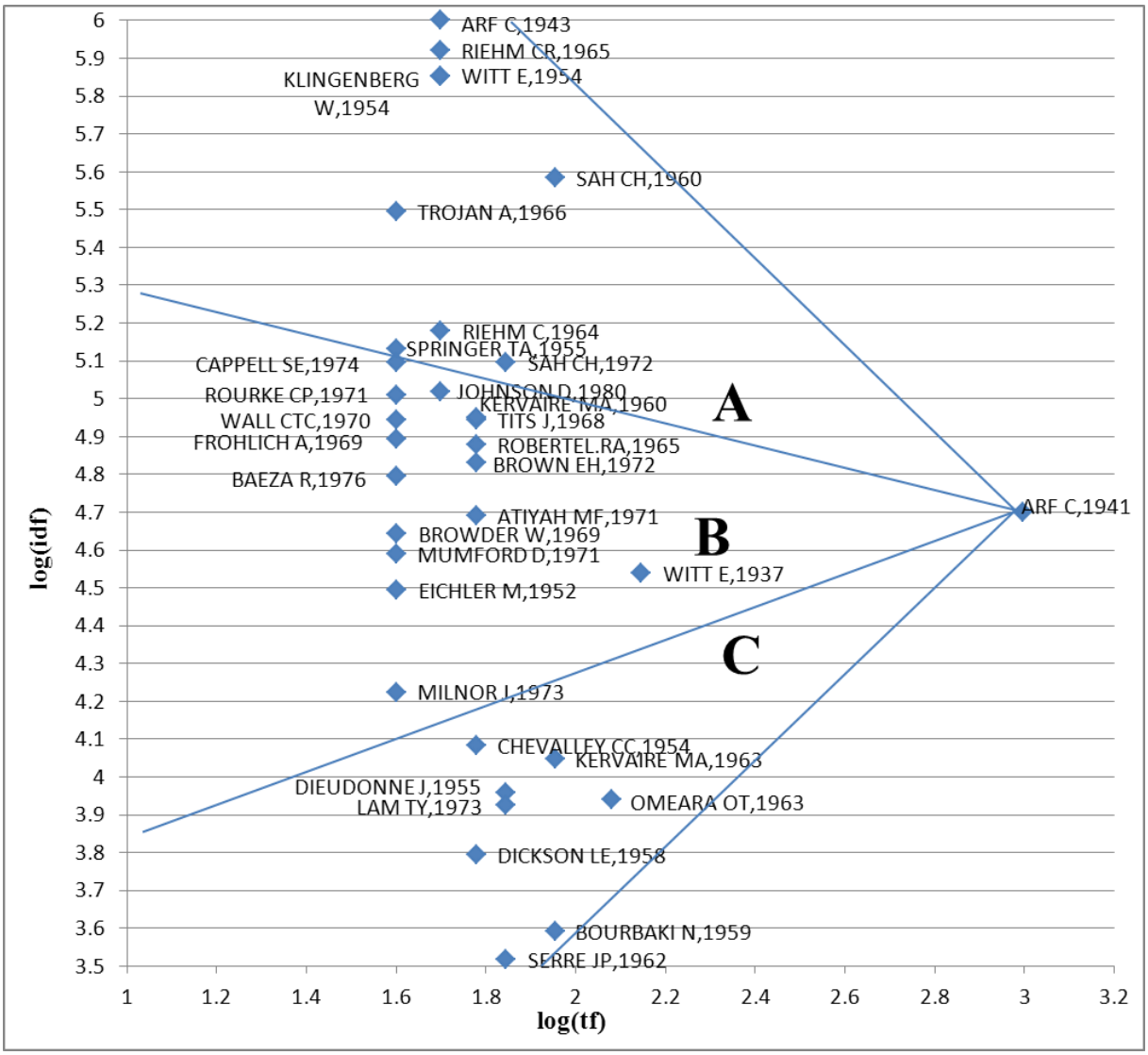

Figure 2. Pennant diagram of items co-cited with Arf's "Untersuchungen über quadratische Formen in Körpern der Charakteristik 2" (Arf, 1941)

We interpret Arf's pennant diagram based on figures given in Table 3, which show how the idf portion of the formula effects the ranking of term distribution. Thomson Reuters' cited reference strings combine both cited works and cited authors, we can use them to make judgments about papers, works and their authors (White, 2007a). Based on these judgments, we draw dividers between sections A, B, and C.

Note that all but one (Cappell SE) authors referred to in Table 3 are placed in section A of Arf's pennant diagram depicted in Figure 2. They have higher tf and idf values in that they are both topically relevant to that of Arf (1941) and their relevance can easily be discerned.

Take, for instance, Arf's own work published in 1943 ("Untersuchungen über quadratische Formen in Körpern der Charakteristik 2. II. Über aritmetische Äquivalenz quadratischer Formen in Potenzreihenkörpern über einem vollkommenen Körper der Charakteristik 2"). It is shown at the top of the pennant diagram because it is a sequel of and complements Arf's original 1941 paper. Papers by Riehm CR, Klingenberg W, Witt E and Sah CH at the top of section A are specifically about quadratic forms over a field of characteristic two, which is exactly the subject of Arf's paper.[17] The topics of papers by the remaining four authors (namely, Trojan A, Riehm C, Springer TA, and Sah C) in section A are all about quadratic forms (but not necessarily quadratic forms over a field of characteristic two) and therefore placed in the relatively lower parts of section A.[18] Note that Chih-Han Sah (entered in two different forms in citation indexes as Sah CH 1960 and Sah C 1972 but corrected in the pennant diagram) is placed higher in the top of the pennant diagram with his specific paper on quadratic forms over a field of characteristic two whereas he is placed relatively lower with his more general paper on symmetric bilinear forms and quadratic forms. So, all nine papers (including Arf's sequel in 1943) in section A are highly relevant to Arf's original paper (Hsia, 1968).

Table 3. Ranks of the first 10 co-cited items with Arf (1941) based on tf and idf

	Freq			*Sector %*			
REFERENCE	*tf*	*df*	*N*	*(tf/df)*100*	*tf*idf*	*log(1+tf)*	*log(5mil/df)*
ARF C,1943	5	5	5E+06	28.316	10.194	1.699	6.000
RIEHM CR,1965	5	6	5E+06	28.694	10.059	1.699	5.921
KLINGENBERG W,1954	5	7	5E+06	29.023	9.946	1.699	5.854
WITT E,1954	5	7	5E+06	29.023	9.946	1.699	5.854
SAH CH,1960	9	13	5E+06	34.991	10.914	1.954	5.585
TROJAN A,1966	4	16	5E+06	29.156	8.803	1.602	5.495
RIEHM C,1964	5	33	5E+06	32.796	8.801	1.699	5.180
SPRINGER TA,1955	4	37	5E+06	31.225	8.220	1.602	5.131
SAH C,1972	7	40	5E+06	36.200	9.404	1.845	5.097
CAPPELL SE,1974	4	40	5E+06	31.432	8.166	1.602	5.097

One can easily see that authors in section A are juniors of Arf (not by age, perhaps, but by the topic of their papers) in that they built on or they did further research on the Arf invariant. Note that Witt whose 1937 paper was developed by Arf (1941) wrote a specific paper on Arf invariant later in 1954.

As we move down from section A to section B, the difference between the subject of Arf's paper and those in section B gets more difficult to discern because papers are no longer specifically on quadratic forms. For instance, Cappell's 1974 paper is not on quadratic forms ("Unitary Nilpotent Groups and Hermitian K-Theory. 1"). We therefore drew the line between section A and B just above Cappell SE 1974 in Figure 2. In general, the relationship between Arf's original paper and those in sector B are not necessarily obvious. Yet, authors

[17] The titles of their papers are: "Integral representations of quadratic forms in characteristic 2" by Riehm CR (1965); "Über die Arfsche Invariante quadratischer Formen mod 2" by Klingenberg W (1954); "Über eine Invariante quadratischer Formen mod 2" by Witt E. (1954); and "Quadratic forms over fields of characteristic-2" by Sah CH (1960).

[18] The titles of their papers are: "Integral extension of isometries of quadratic forms over local fields" by Trojan A (1966); "On integral representations of quadratic forms over local fields" by Riehm C (1964); "Quadratic forms over fields with a discrete valuation" by Springer TA, and "Symmetric bilinear forms and quadratic forms" by Sah C (1972).

in section B can be considered peers of Arf in mathematics and they were usually co-cited with Arf (see Table 2). We have already mentioned Witt E 1937 and Kervaire M 1960 earlier. Many of Arf's peers in section B are considered top mathematicians, some with high index scores,[19] some (just like Arf himself) not represented in indexes at all (e.g., Witt E).
White (2007a, p. 556) considers cited authors in section C as "seniors, culture heroes", cited works as serials, generic titles, and world classics, and cited references as books and classic articles. The contributions of authors in section C of Arf's pennant diagram validate White's prediction in that they consist mostly of classic mathematics texts. *Corps Locaux* by Serre JP (1962), *Algebre* by Bourbaki N (1959), *Linear Groups with an Exposition of the Galois Field Theory* by Dickson LE (1958), *Introduction to Quadratic Forms over Fields* by Lam TY (1973), *Introduction to Quadratic Forms* by O'Meara (1963), *La Géométrie Des Groupes Classiques* by Dieudonne (1955), *Algebraic Theory SPI* by Chevalley C (1954) are all placed at the bottom of Arf's pennant diagram along with a highly cited article, "Groups of Homotopy Spheres: I" by Kervaire MA,1963 (which is cited 449 times as of July 25th, 2011). These authors can also be considered seniors of Arf as they authored highly regarded textbooks in mathematics.

Conclusions and Further Research

Cahit Arf's contributions were not properly listed in citation indexes and his h index score or any other bibliometric indicator cannot therefore be calculated properly. In this study, we used SNA and author co-citation network of Arf to study his overall scientific influence retrospectively. Using White's approach, we drew Arf's pennant diagram on the basis of his author co-citation map to reveal his scientific impact (despite the foggy retrospective data that could be gathered from Thomson Reuters).

Arf's cited references and paper titles or keywords with Arf's contributions to mathematics (e.g., "Arf invariant" or "Arf rings") indicate that Arf is still being cited heavily, despite the fact that his last contribution was in 1960s. This is further confirmed by the results of SNA (1966-2011) as Arf continues to play a prominent role in mathematics. Moreover, Arf's pennant diagram based on his most significant paper (Arf, 1941) clearly shows his overarching influence on generations of mathematicians. Findings obtained through SNA and the pennant diagram seemed to be similar in that some pivotal authors on the co-citation network appeared as peers of Arf in the pennant diagram. However, further work is needed to compare the results of the two methods more comprehensively. As Arf's pennant diagram is based on his single work, we may be missing some crucial authors or works. Further work is also needed to find out if Arf's implicit citations (from titles and keywords of papers) can be incorporated in such analyses using a somewhat different approach. This may provide a fuller picture of Arf's scientific influence.

This study clearly shows that White's (2007a) approach and pennant diagrams can be used to study the impact of authors who are no longer active or their h index scores cannot be calculated on the basis of available data. Once a proper method is developed in the future to incorporate implicit citations to pennant diagrams as co-citations, they can be used to calculate authors' retrospective h index scores. One can conjecture that explicit citations in reference lists of papers would be placed in section A of a pennant diagram, while implicit citations contained in paper titles and keywords be placed in section B and section C, respectively. Further research is needed to validate this conjecture.

[19] For example, Atiyah MF with h index 34, Milnor J with 29, Wall CTC with 25, Cappell SE with 17, Frohlich A with 16, and Kervaire M with 9.

Acknowledgements

We are grateful to Professor Tosun Terzioğlu of Sabancı University for his comments on an earlier draft of this paper.

References

Al, U., Şahiner, M. & Tonta, Y. (2006). Arts and Humanities literature: Bibliometric characteristics of contributions by Turkish authors. *Journal of the American Society for Information Science & Technology*, 57, 1011-1022.

Arf, C. (1941). Untersuchungen über quadratische Formen in Körpern der Charakteristik 2. (Teil I.) *Journal für die Reine und Angewandte Mathematik*, 183, 148-167.

Arf, C. (1948). Une interpretation algébrique de la suite des ordres de multiplicité d'une branche algébrique. *Proceedings of the London Mathematical Society*, Series 2, 50, 256-287.

Arf, C. (1990). *The Collected Papers of Cahit Arf*. İstanbul*:* Türk Matematik Derneği.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57, 359–377.

Chen, C., Song, I.Y., Yuan, X. & Zhang, J. (2008). The thematic and citation landscape of *Data and Knowledge Engineering* (1985-2007). *Data and Knowledge Engineering*, 67, 234-259.

Egghe, L., Guns, R. & Rousseau, R. (2011). Thoughts on uncitedness: Nobel laureates and Fields medalists as case studies. *Journal of the American Society for Information Science and Technology*, 62, 1637–1644.

Garfield, E. (1980). Is information retrieval in the arts and humanities inherently different from that in science? The effect that ISI's citation index for the arts and humanities is expected to have on future scholarship. *Library Quarterly*, 50, 40–57.

Hsia, J.S. (1968). A note on the integral equivalence of vectors in characteristic 2. *Mathematische Annalen*, 179, 63–69.

Ikeda, M.G. (1998). Cahit Arf's contribution to algebraic number theory and related fields, *Turkish Journal of Mathematics*, 22, 1-14.

Lipman, J. (1971). Stable ideals and Arf rings. *American Journal of Mathematics*, 93, 649-685.

McCain, K.W. (1984). Longitudinal author cocitation mapping: the changing structure of macroeconomics. *Journal of the American Society for Information Science*, 35, 351-359.

McCain, K.W. (1986). Cocited author mapping as a valid representation of intellectual structure. *Journal of the American Society for Information Science*, 37, 111-122.

McCain, K.W. (1990). Mapping authors in intellectual space. A technical overview. *Journal of the American Society for Information Science*, 41, 433-443.

McCain, K.W., Verner, J.M., Hislop, G.W, Evanco, W. & Cole, V. (2005). The use of bibliometric and knowledge elicitation techniques to map a knowledge domain: Software Engineering in the 1990s. *Scientometrics*, 65, 131-144

Manning, C.D., & Schütze, H. (2000). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press.

O'Connor, J. J., & Robertson, E.F. (1998). Arf biography, MacTutor History of Mathematics archive. Retrieved, July 18, 2011, from http://www-history.mcs.st-andrews.ac.uk/Biographies/Arf.html.

Otte, E., & Rousseau, R. (2002). Social network analysis: A powerful strategy, also for the information sciences. *Journal of information Science*, 28, 443–455.

Önder, T. (1990). Arf invariant and its applications in topology. In Arf, C. *The Collected Papers of Cahit Arf* (pp. 413-415). İstanbul*:* Türk Matematik Derneği.

Sertöz, S. (1990). On Arf rings. In Arf, C. *The Collected Papers of Cahit Arf* (pp. 416-419). İstanbul*:* Türk Matematik Derneği.

Sertöz, S. (1997). Arf rings and characters. *Note di Matematica*, 14, 251-261.

Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application to retrieval. *Journal of Documentation*, 28, 11-21.

Stern, M. (1983). Characteristics of the literature of literary scholarship. *College & Research Libraries*, 44, 199–209.

Terzioğlu, T. (1998 February). Sunuş yerine (Introduction to the special issue on Cahit Arf). *Bilim ve Teknik*, No. 363. Retrieved, August 7, 2011, from http://www.biltek.tubitak.gov.tr/bdergi/ozel/arf/sunus.html.

Tonta, Y., & Darvish, H.R. (2010). Diffusion of Latent Semantic Analysis as a research tool: A social network analysis approach. *Journal of Informetrics*, 4, 166-174.

White, H.D. (2007a). Combining bibliometrics, information retrieval, and relevance theory. Part 1: First examples of a synthesis. *Journal of the American Society for Information Science and Technology*, 58, 536-559.

White, H.D. (2007b). Combining bibliometrics, information retrieval, and relevance theory. Part 2: Some implications for information science. *Journal of the American Society for Information Science and Technology*, 58, 583-605.

White, H. D. (2009). Pennants for Strindberg and Persson. In *Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at His 60th Birthday*. Special Volume of the *E-Newsletter of the International Society for Scientometrics and Informetrics*, v. 5-S: 71-83.

White, H.D. (2010). Some new tests of relevance theory in information science. *Scientometrics*, 83:653–667.

White, H.D., & Griffith, B.C. (1981). Author co-citation: a literature measure of intellectual structure. *Journal of the American Society for Information Science*, 32,163-171.

White, H.D., & McCain, K.W. (1998). Visualizing a discipline: An author co-citationanalysis of information science, 1972-1995. *Journal of the American Society forInformation Science*, 59, 2146-2155.

Witt, E. (1937). Theorie der quadratischen Formen in beliebigen Körpern. *Journal für die Reine und Angewandte Mathematik*, 176, 31–44.

APPENDIX 1: tf and idf values for authors in Arf's pennant diagram

	weight	log(tf)	log(idf)
	tf*idf	log(1+tf)	log(5mil/df)
ARF C,1941,J REINE ANGEW MATH,V183,P148	14.076	2.996	4.699
SECTOR A			
ARF C,1943,RFSUI A,V8,P297	10.194	1.699	6.000
RIEHM CR,1965,AM J MATH,V87,P32	10.059	1.699	5.921
KLINGENBERG W,1954,J REINE ANGEW MATH,V193,P121	9.946	1.699	5.854
WITT E,1954,J REINE ANGEW MATH,V193,P119	9.946	1.699	5.854
SAH CH,1960,AM J MATH,V82,P812	10.914	1.954	5.585
TROJAN A,1966,CANADIAN J MATH,V18,P920	8.803	1.602	5.495
RIEHM C,1964,AM J MATH,V86,P25	8.801	1.699	5.180
SPRINGER TA,1955,INDAG MATH,V17,P352	8.220	1.602	5.131
SAH C,1972,J ALGEBRA,V20,P144	9.404	1.845	5.097
SECTOR B			
CAPPELL SE,1974,B AM MATH SOC,V80,P1117	8.166	1.602	5.097
JOHNSON D,1980,J LOND MATH SOC,V22,P365	8.525	1.699	5.018
ROURKE CP,1971,ANN MATH,V94,P397	8.024	1.602	5.009
KERVAIRE M,1960,COMMENT MATH HELV,V34,P257	8.803	1.778	4.951
TITS J,1968,INVENT MATH,V5,P19	8.790	1.778	4.943
WALL CTC,1970,PROC CAMB PHILOS S-M,V67,P243	7.919	1.602	4.943
FROHLICH A,1969,J ALGEBRA,V12,P79	7.839	1.602	4.893
ROBERTEL.RA,1965,COMMUN PUR APPL MATH,V18,P543	8.676	1.778	4.879
BROWN EH,1972,ANN MATH,V95,P368	8.588	1.778	4.830
BAEZA R,1976,LECT NOTES MATH,V655,P	7.683	1.602	4.796
ATIYAH MF,1971,ANN SCI ECOLE NORM S,V4,P47	8.340	1.778	4.690
BROWDER W,1969,ANN MATH,V90,P157	7.437	1.602	4.642
MUMFORD D,1971,ANN SCI ECOLE NORM S,V4,P181	7.351	1.602	4.588
WITT E,1937,J REINE ANGEW MATH,V176,P31	9.738	2.146	4.538
EICHLER M,1952,QUADRATISCHE FORMEN,V,P	7.201	1.602	4.495
MILNOR J,1973,SYMMETRIC BILINEAR F,V,P	6.768	1.602	4.225
SECTOR C			
CHEVALLEY CC,1954,ALGEBRAIC THEORY SPI,V,P	7.260	1.778	4.083
KERVAIRE MA,1963,ANN MATH,V77,P504	7.908	1.954	4.047
DIEUDONNE J,1955,GEOMETRIE GROUPES CL,V,P	7.303	1.845	3.958
OMEARA OT,1963,INTRO QUADRATIC FORM,V,P	8.191	2.079	3.939
LAM TY,1973,ALGEBRAIC THEORY QUA,V,P	7.244	1.845	3.926
DICKSON LE,1958,LINEAR GROUPS EXPOSI,V,P	6.749	1.778	3.795
BOURBAKI N,1959,ALGEBRE,pCH9,P	7.021	1.954	3.593
SERRE JP,1962,CORPS LOCAUX,V,P	6.492	1.845	3.518