Notes and Discussion

Chronological Distribution of Information in Historical Texts

JORDAN TABOV
Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria
E-mail: tabov@math.bas.bg

Abstract. In their papers, Kalashnikov *et al.* (1986), Rachev *et al.* (1989) and Fomenko *et al.* (1990) introduced the so-called "volume function" describing the chronological distribution of information in historical texts. Here we give another approach to constructing similar functions.

1. Volume Function

As a general principle, one chronicle cannot describe everything that has happened. Usually the author pays attention to only what is interesting and important to himself and to what he wants to inform the readers about. But such interesting things happen only occasionally, "from time to time", and are separated by "boring periods" in various lengths. That is why in each historical text different years are represented in different ways: much has been written on some of them (which means that many interesting things have happened in those years), less on others, next to nothing on the rest.

So the different time periods are represented by different amounts of information, i.e., the information given by a chronicle is distributed irregularly in the individual time intervals.

It turns out that by following this idea good tools for analysis and comparison of historical texts could be developed, and this has been done in the joint paper of Kalashnikov *et al.* (1986). The method developed by them is based on the use of functions reflecting the distribution of information volume "along the time axis".

Here is the method of obtaining such a function:

We consider a given chronicle, for example *History* of Titus Livius (1989). It covers the interval from 753 BC to 236 BC. We divide this interval to smaller intervals of 20 years each, which seems to be most appropriate in this case. For each interval T_i we "measure" its corresponding information in *History*; for example, we count the number n_i of lines in the book concerning the events that happened in T_i. The number n_i is a simple quantitative parameter of the "volume" of what we know about the period T_i. Using the numbers n_i we construct the "volume function" $L(t)$ in respect of Titus Livius' work, where t denotes the time; t varies from -753 to
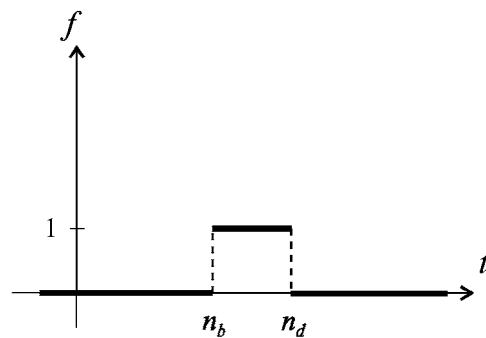
Figure 1. An example of a personal function.

−236. The construction could be described visually in the following manner: in a Cartesian co-ordinate system we graph the time along the x axis and the number of pages along the y axis; more precisely, over each interval T_i we draw a segment parallel to the x axis whose points have ordinates equal to n_i. We consider these segments as a "staircase-like graph" which we "smooth" into a continuous graph by applying standard mathematical methods. As a result we obtain the graph of the "volume function" *vol* $L(t)$.

In Fomenko, 1993 p. 69, the function *vol* $L(t)$ was compared with the volume function *vol* $G(t)$ of the classical *History of Rome in the Middle Ages* by Gregorovius (1900), which describes the period from 300 AD to 816 AD. However, the latter's graph was "slid" leftward along the x-axis by 1053 years, i.e., it was moved back in time by the said amount of years. Figure 9.3–"a" in Fomenko, 1993, p. 96, shows quite eloquently the "similarity" between the both graphs. For conclusions, more details and applications see Fomenko *et al.* (1990).

2. Names as Chronological Information Carriers

We already know the idea of using the volume function to compare historical texts. Now we shall consider a similar function which could be used in a similar way for chronological purposes.

Let T be an arbitrary historical text and let N be some name mentioned in it. Let the dates of birth and death of the person with name N be n_b and n_d respectively. We define the *personal function* (or PF for short) $f_N(t)$ by the formula

$$f_N(t) = \begin{cases} 1, & \text{if } t \in [n_b, n_d] \\ 0, & \text{if } t \notin [n_b, n_d] \end{cases}$$

The graph of $f_N(t)$ is shown in Figure 1.

Now for each name N_i mentioned in T we take the respective PF $f_{N_i}(t)$ and add up all such functions. The result we call Chronological Distribution of Information (CDI) $f_T(t)$ of the text T.
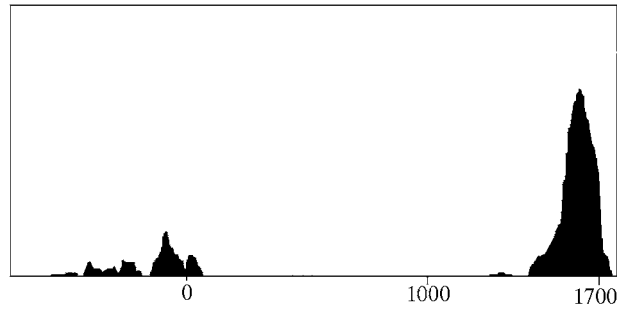
Figure 2. CDI of Lord Chesterfield's *Letters to His Son.*

To put it in a different way,

$$f_T(t) = \Sigma_i f_{N_i}(t). \tag{1}$$

As an indicative example we shall consider the famous work *Letters to His Son* of Lord Chesterfield (1694–1777) (Lord Chesterfield, 1978). In it 151 names of historical persons are mentioned (with different rate frequency). The graph of the respective CDI $f_T(t)$, where T = *Letters to His Son* (the data were extracted from the text by L. Tabova and the names were taken with limited rate frequency – up to 3), is shown in Figure 2. Since we are interested in the approximate graph of $f_T(t)$, the function has been "smoothed".

Letters to His Son could be considered as a manual used by the father to teach his son through historical examples. It contains a number of examples drawn from the life and conduct of famous people (rulers, commanders, poets, artists, noblemen, etc.) and their reactions in various situations. The analysis and interpretation of the results of their behaviour are based on the comments and recommendation of the author to his son and the other readers. Lord Chesterfield's erudition and the moral principles, together with the material selection and the fascinating writing style, made the *Letters* one of the bestsellers of the contemporary and further generations.

Let us suppose that we are 50–60-year-old parents and we would like to write a similar book for our children now, at the very beginning of the 21th century.

Most likely, the major part of our examples would be drawn out of the last 30–40 years; a smaller part would be taken of the mid 20th century, some of the beginning of the century and only occasional facts would be taken of the past ages; perhaps we would use several "ancient" events (just to demonstrate how knowledgeable we are; the ancients' experience would hardly be useful nowadays).

Given the above considerations, the graph of $f_T(t)$ for our book would have – in rough approximation – the shape shown in Figure 3. In particular, it would be "almost zero" outside the 20th century borders.

Comparing this "expected" graph with the graph of Lord Chesterfield's *Letters*, we can see that these two graphs correlate quite well.
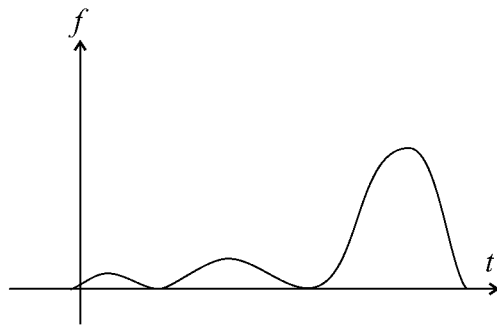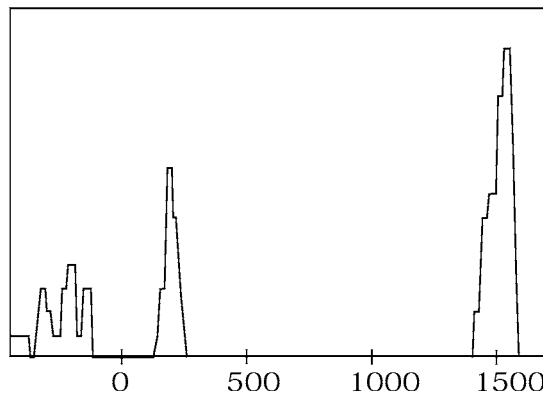
Figure 3. An example of perfect CDI.

Figure 4. CDI of *Il Principe*.

And now let us consider a similar graph (Figure 4) made with respect of the famous book *Il Principe* of N. Machiavelli (Machiavelli, 1991); data were extracted by B. Lazarov. We would like to point out this book is a sort of "textbook" for rulers and diplomats. It contains a number of world history examples used by Machiavelli as a basis to construct his advice – how a ruler should behave and rule and how diplomacy should be transacted.

Comparing the "expected" graph in Figure 3 and the Machiavelli's graph in Figure 4, we find out that they are partially similar. More precisely, this is true for the middle and the right parts of the graph. Meanwhile, the left part of Machiavelli's graph shows a peculiarity in comparison with the "expected" one – there are two "peaks" in the intervals $[-250, -120]$ and $[120, 260]$ which are higher than what we defined to be a normal deviation. They show that:

1. Machiavelli used quite a large amount of information from 3rd to 2nd century BC and from 2nd to 3rd century AD.

   This fact could have different explanations; one of them could be that Machiavelli was very strongly influenced by the cultural and political achievements of the ancient Roman Empire and, as many other representatives of the Italian Renaissance, he might have idealised "ancient Rome". However, if this

is true, then another problem arises: why did Machiavelli not mention anything about the "most brilliant period" (according to the Italian Renaissance makers), from 50 BC to 120 AD – the ages of Caesar and Augustus, Cicero, Claudius, Nero, Trajan and Hadrian?

2. Machiavelli did not use any information from 3rd to 13th century AD.
   That is, Machiavelli did not find anything worthy of mention from the reign of Constantine I "the Great", Justinian, Charlemagne, anything interesting in the Crusades.

What are these peculiarities of Machiavelli's book due to? One possible answer to this question is related to a probable error in the chronology scheme, which "prolongs", replaces and duplicates intervals of the time axes, resulting in deformation of the graph. Such examples are indicative of the necessity of careful exploration and review of CDI.

3. Another Idea of How to Use CDI

Finally, we shall give another idea of how to use CDI to study information sources used in writing a certain historical text T.

Let f_{T_1} be a CDI of another text T_1. If $f_{T_1} \cdot f_T = 0$, then T_1 and T do not contain a relatively considerable amount of general information. In such case T_1 can't be an important information source for T and *vice versa*. Therefore T_1 and T are "independent" of each other, i.e., none of them could be a source for the other one. If we wish to find the information sources within a set $\{C\}$ of other texts, then the above considerations and the comparison of the respective graphs could reduce our search to those texts from $\{C\}$ that are dependent on T. The selection of the latter could be done with the help of a computer.

4. Generalisation

The construction of f_T as described above could be summarised. In our reasoning we got the CDI of a given text T (*Letters to His Son* by Lord Chesterfield in our major example) of the *personal names* mentioned in T. But there is no obstacle to this list of names being extended with other "key" objects, i.e., specific (important) events, e.g., wars, invasions, church councils, etc. For each such object we could consider its relevant function defined in a similar way as the personal function. Actually, in such a way we could expand our concept N_i to a wider range of objects, and keep the same formula for $f_T(t)$ (1) by adding more PFs. We could add the possibility of using "weights" for the separate addenda in this construction and this gives the concept of Generalized Chronological Distribution of Information (GCDI) defined by the formula

$$g_T(t) = \Sigma_i a_i f_{N_i}(t),$$

where a_i, $i = 1, 2, 3, \ldots$ are the coefficients reflecting the "importance" of the respective objects N_i (personal names, wars, church councils, etc). For example, according to Fomenko (Kalashnikov *et al.* (1986), Rachev *et al.* (1989), Fomenko *et al.* (1990), Fomenko (1981)) the "volume function" is a special case of GCDI g_T, in which the chapters (Fomenko, 1981, 1990) correspond to N_i; then the coefficients a_i are proportional to the number of lines in the relevant "chapters" N_i. The selection of a_i itself depends on the purposes of the study where GCDI is applied.

References

Fomenko A. (1981) New Experimental and Statistical Methods of Dating Ancient Events and Application to The Global Chronology of Ancient World. *Preprint Gos. Kom. Telev. Radioveshch.*, 3672(B07201)(9/XI–81), Moscow (in Russian).

Fomenko A. (1990) *Methods of Statistical Analysis of Narrative Texts and Application to Chronology*. Publishing House of Moscow University, Moscow (in Russian).

Fomenko A., Rachev S. (1990) Volume Functions of Historical Texts and The Amplitude Correlation Principle. *Computers and the Humanities*, 24, pp. 187–206.

Fomenko A. (1993) *Critics of Ancient and Middle Ages Traditional Chronology. What Century Is It Now?* Publishing House of the Faculty of Mechanics and Mathematics of the Moscow University, Moscow (in Russian).

Gregorovius F. (1900) *History of Rome in the Middle Ages*. St Petersburg (in Russian).

Kalashnikov V., Rachev S., Fomenko A. (1986) New Methods of Comparing Volume Functions of Historical Texts. In *Proceedings of the Seminar, BNIISI, Problems of Stochastic Models Stability*, Moscow, pp. 33–45 (in Russian).

Lord Chesterfield. (1978) *Letters to His Son. Maxims. Letters*. Nauka, Moscow (in Russian).

Machiavelli N. (1991) *Il Principe*. ESPAS, Sofia (in Bulgarian).

Rachev S., Fomenko A. (1989) Volume Functions of Historical Texts and the Amplitude Correlation Principle. In *Methods of Studying Historical Sources about Public Thoughts during Feudalism, Collection of papers*, Institute of USSR History (USSR Academy of Sciences), Moscow, pp. 161–180 (in Russian).

Titus Livius. (1989) *History of Rome*. Volumes I–III, Nauka, Moscow (in Russian).