Overview of Network Information by Using Anchored Maps

Kazuo Misue

Department of Computer Science, University of Tsukuba,
1-1-1 Tennoudai, Tsukuba, 305-8573 Japan
misue@cs.tsukuba.ac.jp

Abstract. Spring embedding is widely used to draw up networks. However, the readability of network diagrams drawn by using spring embedding is still not good enough. This paper proposes the use of "anchored maps," in which some of the nodes are fixed as anchors to improve the readability of network diagrams. First, the readability of network diagrams is discussed and the anchored maps are proposed. Then, an overview of the method the author developed for drawing anchored maps is explained. Next, examples of network diagrams that were created from a real sales database are used to prove the effectiveness of anchored maps.

Keywords: graph drawing, anchored map, visualization, knowledge mining.

1 Introduction

The purpose of this research is to develop visualization techniques to give overviews of networks that are highly readable for mining knowledge. The purpose of giving these overviews is to support the first movement to extract useful knowledge from network information. When we process large-scale numeric data, we usually draw graphs (e.g., scatter charts and histograms) to grasp the overall tendency of the data. When we process network information, we notice that we do not have enough techniques and tools to manipulate it. Techniques and tools for this type of manipulation are currently being developed.

The author proposes the introduction of dependable viewpoints from observations into network diagrams to improve the readability of the diagrams. It is possible to visualize network information as diagrams. However, it is often difficult to get any useful knowledge from these types of diagrams. Introducing viewpoints into network diagrams is expected to enable observers to adequately read useful information from them.

As an approach to introducing viewpoints on diagrams drawn by using spring embedding, which is often used to layout networks, we chose to fix some nodes at predetermined positions. The fixed nodes act like coordinate systems and perform functions as viewpoints. The author named this style of drawing "anchored map." Similar representation techniques are also used in some other tools, for example, Visual Who [2], which is a visualization tool for communities and SQWID [4], which is a WWW search tool.

B. Apolloni et al. (Eds.): KES 2007/ WIRN 2007, Part II, LNAI 4693, pp. 1269–1276, 2007.

In the beginning, an anchored map has just three anchors. However, four or more viewpoints are often necessary, so general techniques for drawing anchored maps with four or more fixed nodes are necessary.

2 Readability of Network Diagrams

Drawing network diagrams (represent networks visually) and enabling to ability to read some of the meanings from the diagrams are not equal. If we expect the meanings of network diagrams to be readable, conventional techniques to draw networks (graph drawing techniques) cannot provide diagrams with enough readability.

2.1 Readability of the Spring Embedding

It has been a long time since we have seen the tools and applications that draw network diagrams by using spring embedding [3]. The author also has experience in developing the programs of spring embedding. The purpose of the programs is not just for drawing undirected graphs, but also for the representation of the relationships between concepts as network diagrams. The programs seem to work well at making the layouts of undirected graphs. However, when we tried to interpret the meaning in the application area of the networks, we often had trouble understanding what the diagrams expressed as a whole.

It is clear that the network diagrams represent the relationships between the nodes in each local segment. However, if we want to express the local relationships among the nodes, we do not need to draw network diagrams; tables and list expressions are also useful and are good enough for this purpose. One of the most important features of diagrams, especially network diagrams, is that we can look at the overview of the whole structure of the network. Drawing network diagrams provides overviews of the whole structure. However, most of us, including this author, are not satisfied with just these diagrams. The spring embedding enabled viewers to see an overview of the network information, but didn’t enable them to read the meanings of the network information.

2.2 Introduction of Viewpoints into Network Diagrams

One of the factors of the low readability in network diagrams drawn using spring embedding is the lack of coordinate systems. Most layout techniques for directed graphs use the features of directed graphs; examples are upward/downward drawings [1] and hierarchical drawing [8]. These techniques orient the directions of the edges in one common direction, so they can visually express global direction. By using such layout techniques, we can grasp the overall structure of the networks and the position of each element in the global structure, even if we do not know the meanings of the nodes and edges.

Network diagrams drawn by using spring embedding do not have coordinate systems. This means that the diagrams do not give viewpoints to read the meanings of the network. To provide diagrams in which the meanings of the networks are highly readable it is critical to provide coordinate systems or viewpoints to the readers.

We developed a method to introduce the proxies of coordinate systems into undirected graphs; the method fixes three nodes chosen by the observers on the three vertices of a triangle [7]. The method fixes the three nodes (called “anchors”), which are the viewpoints, and arranges the other nodes (unfixed nodes, called “free nodes”) by using spring embedding. Each unfixed node is placed in a stable position in the energy of springs through the push-pull processes by the anchors. When we consider the anchors as viewpoints, we can see the position of the free nodes as expressions of the relations with viewpoints.

2.3 Problem with Drawing Anchored Maps

There is a reason why only three anchors are used; it is to keep the anchors symmetrical. When three anchors are used, the network diagrams are independent to the arrangement of the anchors; the diagrams with any anchor arrangements are essentially the same. This means that we do not have to worry about the arrangement of the anchors.

However, the number of nodes to be used as viewpoints is not always three. When the number of anchors used is four or more, the arrangements of the anchors are not symmetrical. So we have to take the arrangement of the anchors into consideration.

We illustrate some of the problems that occurred when four or more anchors were used in Fig. 1. Assume that four anchors (a1, a2, a3, and a4) are arranged on four vertices of a square clockwise in this order. In addition, assume that two free nodes (f1 and f2) are related to two anchors with almost the same degrees. This means that the two free nodes are connected to the anchors by springs with almost the same lengths and strengths. If a free node has relations with a1 and a2, the node is arranged on the edge of the square (see f1 in Fig. 1 (a)). However, if it has relations with a1 and a3, the node is arranged near the center of the square (see f2). The difference in the two positions cannot be neglected. Additionally, when another free node has relations to a2 and a4, the free node is also arranged near the center of the square (see f3 in Fig. 1 (b)). In this case, although f2 and f3 are completely different from each other in relation to the anchors, they are arranged at close places.

It is necessary to choose the appropriate arrangement of anchors to avoid this type of visual misunderstanding.

Fig. 1. Problems with four anchors

By the way, anchors might have orders by their meanings. In such cases, the anchors can be arranged in the orders, even when four or more anchors are used. For example, when we choose the days of the week as anchors, we may arrange the anchors in the order of “Sunday,” “Monday,” …, “Saturday.”

3 Anchored Maps Drawing Method

When four or more anchors are used and the anchors do not have orders, we have to adequately decide the orders of the anchors. This problem cannot be solved by simple extending the spring embedding. We reorganized the drawing method of the anchored maps to solve the problem. A detailed description of the method is shown in [6].

3.1 Drawing Objects

We assume that the drawing objects are bipartite graphs. A bipartite graph is a graph where the set of nodes can be divided into two disjointed sets, such that no edge has both end-points in the same set. One of the two node sets is assumed to be a set of “anchors,” and the other is assumed to be a set of “free nodes.”

3.2 Drawing Conventions

The drawing convention of the anchored maps is written as follows:

(Arrangement convention for the nodes)
- Compound coordinate system: anchors are discretely arranged on a circumference, while the free nodes are freely arranged.

(Routing convention for the edges)
- Straight-line wiring: a straight line connects adjacent nodes.
- The edges are independent of the coordinate system.

3.3 Drawing Rules

We employed the following drawing rules for our anchored map drawing method.

(R1) The nodes are mutually separated more than the lowest distance.
(R2) The nodes connected to each other are laid out as closely as possible (minimizing the total length of the edges).
(R3) The number of crossings among the edges is minimized.
(R4) The anchors connected to common free nodes are laid out as closely as possible.
(R5) The free nodes connected to common anchors are laid out as closely as possible. (Free nodes belonging to the same cluster are laid out as closely as possible; free nodes that do not belong to the same cluster are not closely laid out).

The drawing rules (R1), (R2), and (R3) are also employed in many other drawing methods. We also judged them to be important in consideration of the adaptability to the purpose of drawing anchored maps. Rules (R4) and (R5) are specific to anchored maps. A formal description of these rules is shown in [6].

3.4 Drawing Procedure of Anchored Maps

An anchored map is laid out using the following two steps:

(Step 1) Arrange anchors on the circumference at equal intervals. The size (i.e., radius) of the circumference is decided according to the size of the drawing area (i.e., window), and the order of the anchors on the circumference is decided.

(Step 2) Fix the anchors, and arrange the free nodes at the positions in which their relationships to the anchors are appropriately expressed.

The length of the circumference doesn't influence the quality of the layout; it influences only the size of the drawing. On the other hand, the order of the anchors largely influences the quality of the layout. We do not need to worry about the routing of the edges since the edges are drawn as straight-line segments.

4 Overview of Network Information

An example of the overview of the network information extracted from an actual database to explain the effectiveness of anchored maps is presented in this section.

4.1 Database

A sales database was exploited for experimentation purposes. One transaction is represented as one record, and each record has a purchase date, time, customer name, item name, item category, and price. We could not get a database from a real shop, so we used a sales database managed in our laboratory. It is not a database of a real shop, but the transactions are real. We started to use the database in June 2006, and 2652 purchases were recorded for half a year.

4.2 Overview Using Conventional Means

Various relations can be extracted from the database. In this paper, we focus on the relations between items and time zones (every hour; 00 ~ 23 hrs.) when the items were sold. We removed the items that were sold three times or less in any time zone to avoid noise. What remained was 37 kinds of items, which were sold four or more times in a given time zone.

Table 1 shows the relations between the items and time zones. Each row corresponds to an item, and each column corresponds to a time zone. Each cell represents the number of items sold in a time zone. Often the table is shown in a histogram. This type of representation style is widely used in the early stages of a statistical analysis. By using these representations, we might answer some pointed questions, such as what is the item sold most in a certain time zone. We can say that this representation is suitable for reading the relations between (items, time-zone) and the number of items sold.

Table 1. Overview by using a table

ITEMS	00h	01h	02h	03h	04h	05h	06h	07h	08h	09h	10h	11h	12h	13h	14h	15h	16h	17h	18h	19h	20h	21h	22h
午後の紅茶 ミルクティ	1	5	3	4	3	3	1	0	0	23	5	28	21	35	16	15	19	10	11	4	7	3	6
うまい棒めんたい味	11	8	8	1	6	1	5	0	0	1	1	0	1	2	5	3	9	7	8	5	7	9	5
充実野菜 (930g PET)	5	12	7	2	1	6	2	0	0	3	1	1	4	3	3	3	4	2	5	7	7	6	16
WONDA?ーニングショ	3	3	7	5	3	3	0	0	0	6	2	2	2	10	6	11	5	2	0	8	0	5	9
マジックフレークピーナ	6	0	4	2	1	0	0	0	0	2	2	2	3	8	4	6	21	12	5	6	0	0	3
うまい棒サラミ味	7	9	7	1	2	0	0	0	0	0	0	0	0	2	1	4	5	2	9	10	5	6	7
リポビタンD	3	8	7	2	4	1	2	0	0	3	2	2	6	3	2	6	4	4	3	3	0	2	6
うまい棒コーンポタージ	13	7	8	2	1	1	0	0	0	1	0	0	1	0	2	0	2	1	4	2	4	2	2
うまい棒チーズ味	14	5	3	7	1	0	0	0	0	0	0	0	0	0	2	0	5	3	4	2	0	3	3
卵	1	0	1	4	0	1	0	0	0	6	2	0	6	3	0	0	3	2	3	5	6	2	4
オロナミンC	3	6	5	4	0	2	0	0	0	3	1	1	2	1	0	1	0	2	0	0	2	2	4
マルシンフレッシュソー	1	2	4	1	1	0	0	0	0	0	0	0	1	7	4	1	1	3	7	3	2	1	0
天然ミネラル麦茶	2	0	0	1	0	0	0	0	0	4	3	1	2	2	3	2	7	2	1	2	2	2	0
伊藤園ウーロン茶	2	2	0	0	1	0	0	0	0	2	5	4	9	5	0	3	1	0	0	1	0	1	0
トデカミンゴールド	0	3	2	0	1	0	2	0	0	1	2	0	3	4	1	5	2	2	2	0	3	0	1
明治ミルクチョコレート	4	4	1	1	2	0	0	0	0	0	1	3	1	1	0	3	1	1	3	0	0	1	3
うまい棒テリヤキバー	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	0	0	6	2
コカコーラ	0	0	1	1	1	1	0	0	0	0	0	0	2	3	5	2	3	0	0	1	1	2	0
うまい棒とんかつソー	2	4	1	0	1	2	0	0	0	0	0	0	0	0	0	8	3	0	1	0	0	1	3
野菜生活100(?)	0	0	0	2	1	1	1	0	0	1	0	4	1	0	0	2	1	0	2	1	3	2	1
ボンカレーゴールド21	2	0	0	0	0	2	0	0	0	0	1	2	3	0	1	0	0	0	4	4	1	2	1
おーいお茶	0	1	1	1	0	0	0	0	0	2	0	0	5	2	4	2	0	3	1	0	1	0	1
FIRE???ーワン	1	2	1	2	1	1	0	0	0	0	0	1	0	2	2	1	4	0	1	0	1	2	2
生茶	0	2	1	0	0	0	0	0	0	0	0	2	3	4	2	3	0	1	4	0	1	1	0
KS??????	0	0	2	0	0	0	0	0	0	1	1	1	6	2	0	1	0	0	2	1	2	1	1
QOO?????ーグルト味	0	0	4	0	0	0	1	0	0	1	0	0	1	5	1	1	1	0	1	1	1	1	2
おーいお茶濃い味	1	3	0	1	0	0	0	0	0	2	4	1	2	1	3	1	0	1	0	0	1	0	0
日清焼きそばU.F.O.	1	0	0	0	0	0	0	0	0	0	0	1	3	4	2	0	0	1	1	2	1	2	1
カロリーメイトブロック	2	3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	5	1	1	0	1	0
うまい棒やさいサラダ味	0	3	0	0	0	2	2	0	0	0	0	0	0	0	0	0	3	4	1	0	2	0	0
カントリーマアム	4	0	2	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1	0	0	2	2	0
C.C. Lemon	0	0	0	0	0	0	0	0	0	2	0	1	1	0	2	0	4	0	0	1	1	0	1
東京浅草ソース焼きそ	0	0	1	1	0	0	0	0	0	0	1	1	0	4	1	0	0	0	0	0	4	0	0
わさびのり太郎	1	3	4	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
ホームパイ	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	3	0	0
GokuriMiracle of Grap	1	1	0	0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	1	0	1	1	0
昭和スパゲッティゆで	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	1	0	0

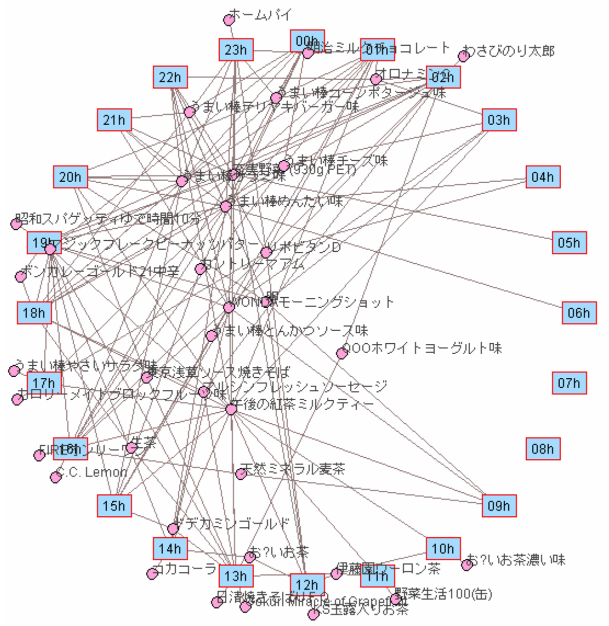
4.3 Overview by Using Anchored Maps

Figure 2 (a) shows the networks that represent the relations between the items and time zones when the items were bought. The time zones are placed as the 24 anchors. The time zones have a cyclic order, so they were placed in their proper order. From this map, we quickly find that the activity of the students is very low in the morning. Of course, if we carefully read Table 1, we can also find this information. This figure also intuitively informs us of this fact. Furthermore, we find that food such as the curries and spaghettis were bought in specific time zones, while canned coffee and sweets did not appear to depend on any specific time zone.

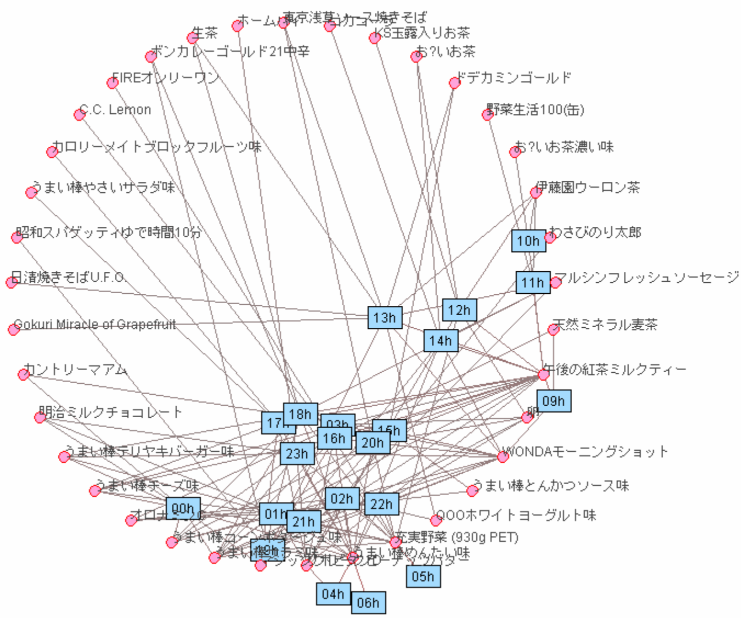
In Fig. 2 (b), the items are placed as the anchors. Since the items have no proper orders, we used the technique described in [6] to decide the order of the anchors. We see the following clusters in the map.

(1) Early morning 1: 04, 05, and 06 hrs.
(2) Early morning 2: 07 and 08 hrs. (no nodes = no activities)
(3) Morning: 09, 10, and 11 hrs.
(4) Lunch time: 12, 13, and 14 hrs.
(5) Evening ~ night: 15 ~ 03 hrs.

It might be possible to find these types of clusters without an overview of the networks using the clustering techniques that already exist. However, the advantage with the network information overview is that we can see the grounds and the background that compose the clusters as the connecting lines between the nodes. By using such maps, we can roughly observe the existence of the clusters from the beginning. Then, we better understand the reason why these clusters exist by touching on a virtual physical model of the relations.



(a) 24 time zones are placed as anchors



(b) 37 items are placed as anchors

Fig. 2. Overview by using anchored maps

5 Conclusion

This paper described the problem with the lack of coordinate systems in topological diagrams, which is the low readability of the diagrams. A drawing style called "anchored maps" was proposed to counteract this problem. In the anchored maps, some nodes, called "anchors", are fixed as viewpoints and thus effectively act like coordinate systems. The author developed a method to draw anchored maps with four or more anchors. The anchored maps are useful at the early analysis stage of network information.

There are several topics remaining as future works; one of them is the conceptual rearrangement of the "readability" of networks. Usually the readability of networks or graphs is described as the drawing conventions and drawing rules. The introduction of the viewpoints proposed in this paper might help to achieve a higher level of conceptual understanding than the drawing conventions and rules. These types of concepts should be organized as the aesthetic criteria for understandability.

References

1. Di Battista, G., Tamassia, R.: Algorithms for plane representations of acyclic digraphs. Theoretical Computer Science 61(2-3), 175–198 (1988)
2. Donath, J.S.: Visual Who: Animating the affinities and activities of an electronic community. In: ACM Multimedia 95 - Electronic Proceedings, ACM Press, New York (1995)
3. Eades, P.: A Heuristic for Graph Drawing. Congressus Numerantium 42, 149–160 (1984)
4. McCrickard, S.D., Kehoe, C.: Visualizing Search Results using SQWID. In: WWW6. Proceedings of the sixth International World Wide Web Conference (1997)
5. Misue, K.: Drawing Bipartite Graphs as Anchored Maps. In: APVIS 2006. Proceedings of Asia-Pacific Symposium on Information Visualization, pp. 169–177 (2006)
6. Misue, K.: Anchored Maps: Visualization Techniques for Drawing Bipartite Graphs. In: Proc. of HCI International (2007) (to appear)
7. Misue, K., Watanabe, I.: Visualization of Keyword Association for Text Mining. In: IPJS SIG Technical Report 1999-FI-55-8, Information Processing Society of Japan, pp. 65–72 (1999) (in Japanese)
8. Sugiyama, K., Tagawa, S., Toda, M.: Methods for Visual Understanding of Hierarchical Systems. IEEE Trans. SMC 11(2), 109–125 (1981)