

WiFIsViz: Effective Visualization of Frequent Itemsets

Carson Kai-Sang Leung* Pourang P. Irani Christopher L. Carmichael
The University of Manitoba, Winnipeg, MB, Canada
 {kleung, irani, umcarmi1}@cs.umanitoba.ca

Abstract

Frequent itemset mining plays an essential role in the mining of many different patterns. Most existing frequent itemset mining algorithms return the mined results—namely, frequent itemsets—in the form of textual lists. However, the use of visual representation can enhance the user understanding of the inherent relations in a collection of frequent itemsets. In this paper, we propose an effective visualizer, called WiFIsViz, to display the mined frequent itemsets. WiFIsViz provides users with an overview and details about the itemsets. Moreover, this visualizer is also equipped with several interactive features for effective visualization of the frequent itemsets mined from various real-life applications.

1. Introduction

Frequent itemset mining [1, 10, 12, 13, 14] aims to search for implicit, previously unknown and potentially useful information—in the form of frequently occurring patterns—that might be embedded in data. It is in demand in many real-life applications, and it also plays an essential role in the mining of many patterns other than frequent itemsets (e.g., maximal itemsets, closed itemsets, association rules, correlations, sequences, episodes). This explains why we have witnessed numerous studies in the area of frequent itemset mining over the last decade. However, most of these studies return a collection of mined frequent itemsets in *textual form* (e.g., a very long list of frequent itemsets). Consequently, users may not easily see the useful information that is embedded in the data.

The common cliché that “a picture is worth a thousand words” advocates the use of visual representation and visualization techniques [7, 8, 18]. Showing a collection of frequent itemsets in *graphical form* can help users understand the nature of the information and show the relations embedded in the data. Hence, to assist users in gaining insight into massive amounts of data or information, researchers have developed various systems for visualizing data [4, 16] and for visualizing the mining results (e.g., clusters [9], decision trees [2], social networks [3], association rules [5, 6, 17]).

Recently, we proposed a frequent itemset visualizer called FIsViz [11]. It represents frequent k -itemsets as poly-lines connecting k nodes (where each node represents an

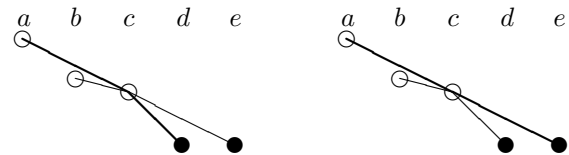


Figure 1. Polylines showing $\{a, c, d\}$ & $\{b, c, e\}$ vs. polylines showing $\{a, c, e\}$ & $\{b, c, d\}$.

item in the k -itemset) in a two-dimensional space. The locations of the polylines indicate the frequencies (or support) of the frequent itemsets. As a result, FIsViz enables users to visualize the mining results—specifically, frequent itemsets—for many real-life applications. However, in some other applications (especially, when the number of frequent itemsets is huge), users may require more efforts to be able to visualize frequent itemsets. The problem is caused by the use of polylines for representing frequent itemsets. To elaborate, the polylines can be bent and/or can cross over each other. This makes it difficult to distinguish one polyline (representing an itemset) from another. For example, in Figure 1, how could one easily distinguish itemsets $\{a, c, d\}$ & $\{b, c, e\}$ from itemsets $\{a, c, e\}$ & $\{b, c, d\}$?

Hence, some natural questions to ask are: Can we use an alternative representation for frequent itemsets to *minimize the polyline crossings*? Can we design a *scalable* system that helps users visualize frequent itemsets effectively, even when the number of frequent itemsets is huge?

Here, we present a visualizer that avoids polyline crossings. Our **key contribution** of this IEEE ICDM 2008 paper is a novel, interactive and scalable system—called WiFIsViz—which uses a wiring-type diagram (or more precisely, orthogonal graphs) for visualizing frequent itemsets. This visualizer provides users with clear and explicit depictions about frequent itemsets that are embedded in the data of interest. Hence, WiFIsViz enables users to infer patterns at a glance and answers many questions in various real-life applications. It also provides users with interactive features for constrained and/or interactive mining. Moreover, with WiFIsViz, users can efficiently find other patterns such as closed or maximal itemsets.

This paper is organized as follows. In Sections 2 and 3, we introduce our WiFIsViz and describe its design as well as its interactive features. Section 4 shows evaluation results. Finally, conclusions are presented in Section 5.

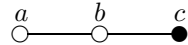
*Corresponding author: C.K.-S. Leung

2. Our proposed frequent itemset visualizer

In this section, we show how wiring-type diagrams (or orthogonal graphs) are used to represent frequent itemsets in our proposed frequent itemset visualizer called *WiFiViz*. By using wiring-type diagrams, WiFiViz minimizes bends and avoids the crossings of the polylines. In general, WiFiViz consists of two views: an *overview* (Section 2.1) on the left and a *detailed view* (Section 2.2) on the right.

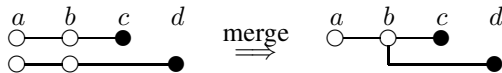
2.1. WiFiViz: the overview

In the overview of our proposed WiFiViz, we show frequent k -itemsets in a two-dimensional space, where the x -axis shows n domain items. Without loss of generality, these domain items are arranged according to some order \mathcal{R} —say, non-ascending frequency order or lexicographical order—in the x -direction. The y -axis shows the frequencies of itemsets. Inspired by wiring-type diagrams, we represent frequent k -itemsets as “wires”. More precisely, a k -itemset X is represented by a horizontal line connecting k nodes (represented by k circles), where each node represents an item in X . For example, the 3-itemset $\{a, b, c\}$ is represented by a horizontal line connecting three circles (where each circle represents an item), as follows:



Note that only one of the three circles is filled. The filled circle represents the last item (according to the item order \mathcal{R}) in the itemset. In WiFiViz, we indicate the frequency of any k -itemset by the y -value of its last item (i.e., the y -position of rightmost filled circle in the horizontal line representing the itemset). A singleton itemset (i.e., 1-itemset), which is a special case of a k -itemset, can be represented by a filled circle (or by a filled square for user convenience—see Property P2 below).

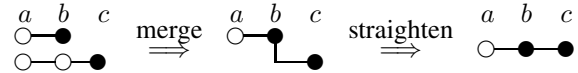
With this representation of itemsets and their frequencies, it is not uncommon that multiple itemsets may have the same frequency. Among them, some may even share the same prefix. To allow users to easily visualize these frequent itemsets, WiFiViz reduces the number of circles (filled or hollow) to be displayed by applying *path-sharing compression*. To elaborate, if two itemsets X and Y are of the same frequency and share the same prefix, then their common prefix can be merged. The suffixes of X and Y are then branching out from the last item in the common prefix. For example, if itemsets $\{a, b, c\}$ and $\{a, b, d\}$ (which share the same prefix $\{a, b\}$) are of the same frequency, WiFiViz represents them as follows:



Here, $\{c\}$ and $\{d\}$ are two branches of the common prefix $\{a, b\}$.

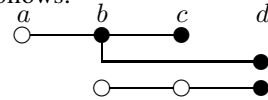
A special case of this path-sharing compression occurs when X is a prefix of Y . In this case, the two lines representing the two itemsets X and Y would be merged (and

straightened) into one horizontal line. For example, for itemsets $\{a, b\}$ and $\{a, b, c\}$, the former is a prefix of the latter. Hence, WiFiViz merges (and straightens) the lines representing these two itemsets to form the following:

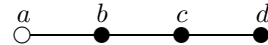


Here, the filled circle c indicates the last item in itemset $\{a, b, c\}$, whereas the filled circle b indicates the last item in the prefix $\{a, b\}$. Note that this compression helps reduce the number of horizontal lines to be drawn (i.e., reduce the required amount of vertical space to display all frequent itemsets) without losing any information.

When dealing with a huge number of frequent itemsets, multiple itemsets may have the same frequency. Some of them may not share the same prefix. Hence, to further reduce the required amount of vertical space, WiFiViz applies another compression technique—*projection*. To elaborate, if m itemsets (represented by $\leq m$ horizontal lines) have the same frequency, WiFiViz projects their corresponding lines onto one horizontal line. For instance, let us suppose itemsets $\{a, b\}$, $\{a, b, c\}$, $\{a, b, d\}$ and $\{b, c, d\}$ to be of the same frequency. Then, these itemsets can be represented as follows:



In WiFiViz, these four itemsets (represented by three lines) are projected (or compressed) onto a horizontal line in the overview, as shown below:



By doing so, each existing frequency value (or range of frequency value, if necessary) would be represented by *one and only one* horizontal line. This greatly reduces the number of horizontal lines to be drawn. However, the tradeoff is that we *may* lose information in some situations. For example, the above horizontal lines may be the result of applying projection to lines representing itemsets $\{a, b\}$, $\{a, c\}$ and $\{a, d\}$. This explains why, in addition to the overview, WiFiViz also provides users with the detailed view (Section 2.2).

It is important to note, in some other situations, users are able to precisely and easily visualize frequent itemsets without using the detailed view (see Properties P5 and P6 below) or at least to tell which items that are present in (or absent from) the itemsets (see Properties P7 and P8).

With WiFiViz, one can observe and recognize the following *properties of the overview*:

P1. For each domain item, WiFiViz gives information about its frequency distribution. Specifically, the overview tells how many distinct frequency values an item takes on. For example, by counting the number of nodes (filled and hollow circles as well as filled

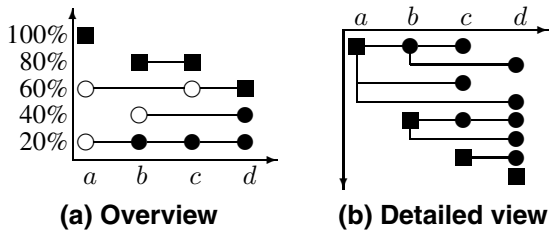


Figure 2. The two views in WiFIsViz.

squares) for item b in Figure 2(a), we know that any itemsets containing b have one of the three frequencies: 80%, 40% or 20%.

- P2. For user convenience (e.g., ease of item recognition), WiFIsViz represents each singleton itemset α by a *filled square* so that the frequency of α is indicated by the y -position of the square. Alternatively, WiFIsViz can also represent α by a *filled circle* so that the frequency of α is indicated by the highest y -position of all filled circles for α . For example, in Figure 2(a), the highest y -position of all filled nodes for domain item b is 80%. This is the frequency of b .
- P3. By default, WiFIsViz arranges domain items in non-ascending frequency order. As a result, the most frequently occurring item (i.e., the item with the highest frequency) appears on the left side of the overview and the least frequently occurring one appears on the right side. (It is important to note that users are not confined to this order. They can choose any order \mathcal{R} to arrange domain items.)
- P4. WiFIsViz provides users with an intuitive overview of frequency distribution of all domain items. For example, Figure 2(a) shows that all domain items have frequencies $\geq 60\%$ (as indicated by y -positions of filled squares or by highest y -positions of filled nodes).
- P5. If a horizontal line involves *only two* items (say, b and c), then users can get complete information about this 2-itemset $\{b, c\}$ without using the detailed view. The frequency of this 2-itemset is clearly indicated by the y -position of the last item c .
- Moreover, if such a horizontal line (for frequency $y\%$) starts with a filled node, then this line gives additional frequency information that $\text{sup}(\{b\}) = \text{sup}(\{b, c\}) = y\%$.
- P6. If a horizontal line (connecting two or more items) starts with a hollow circle (say, representing item a) and is followed by a filled node (say, representing item b), then users can reveal—without requiring the detailed view—that 2-itemset $\{a, b\}$ exists with that frequency.
- P7. WiFIsViz provides information about the *presence* of items in some itemsets. For example, when a node (filled or hollow) representing item c appears on the

horizontal line for frequency 60%, we know that c is contained in at least one itemset having frequency 60%.

- P8. WiFIsViz also provides information about the *absence* of items from any itemsets. For example, when item b does not appear in the horizontal line for frequency 60%, we know that b is guaranteed not to appear in any itemsets of that frequency.
- P9. The frequency of any subset of a k -itemset X is guaranteed to be higher than or equal to that of X . Hence, the filled square representing any singleton subset of X , as well as the horizontal line representing any non-singleton subset of X , is guaranteed to appear on or above the horizontal line representing X .
- P10. The frequency of any superset of a k -itemset X is guaranteed to be lower than or equal to that of X . Hence, the horizontal line representing any superset of X is guaranteed to appear on or below the horizontal line representing X .

2.2. WiFIsViz: the detailed view

In the previous section, we showed the overview of our proposed WiFIsViz. In this section, let us present the *detailed view* of WiFIsViz. Recall that the overview aims to provide users with a quick insight about the distribution of itemsets. Using different compressions (e.g., path-sharing, projection), we can reduce the number of nodes and/or horizontal lines that need to be displayed. While users are able to visualize frequent itemsets using only the overview in some situations, there are other situations in which the detailed view is needed.

In the detailed view of our proposed WiFIsViz, frequent k -itemsets are also represented in a two-dimensional space. As in the overview, the n domain items are arranged according to some order \mathcal{R} (e.g., non-ascending frequency order) in the x -direction. Unlike the overview, frequent itemsets are *not* arranged according to their frequencies in the y -direction. Instead, similar itemsets (i.e., those that share common prefixes, regardless of their frequencies) are grouped together in the detailed view. We also apply path-sharing compression, but *not* projection compression, to reduce the amount of vertical space required to display all these frequent itemsets. See Figure 2(b) for the detailed view (cf. Figure 2(a), which shows the same collection of frequent itemsets in the overview). With WiFIsViz, one can observe and recognize the following *properties of the detailed view*:

- P11. For each horizontal line in the detailed view, if a node (representing an itemset X) is filled, all the nodes to its left (representing all prefixes of X) must be filled. This is due to the Apriori property that all prefixes (subsets) of a frequent itemset are guaranteed to be frequent.

P12. By counting the number of nodes on the horizontal line (from the leftmost node to a filled node) in the detailed view, we can obtain the cardinality k of the k -itemset represented by that line (e.g., cardinality of $\{a, b, c\}$ is 3 because there are 3 nodes from the leftmost square to the second filled circle). While users can count the number of nodes to determine the cardinality of an itemset, WiFIsViz provides an interactive feature called *query on cardinality* for user convenience. See Section 3.

In WiFIsViz, the detailed view and the overview complement each other. The two views are linked so that users can highlight one or more horizontal lines in one view (i.e., “brushing”) and the corresponding horizontal lines are also highlighted in the other view. One can observe and recognize the following *properties of the two views*:

P13. When we select a horizontal line of frequency $y\%$ in the overview, WiFIsViz highlights all itemsets of such frequency in the detailed view (i.e., finds out exactly what itemsets were projected onto the line of frequency $y\%$). For example, when we select the horizontal line of frequency 60% in the overview, WiFIsViz highlights in the detailed view itemsets $\{a, d\}$, $\{c, d\}$ and $\{d\}$ (all with frequency 60%).

P14. Conversely, when we select an itemset X (or several itemsets) in the detailed view, WiFIsViz highlights the corresponding horizontal line in the overview. This reveals the frequency of X . For example, when we select $\{a, b, c\}$ in the detailed view, WiFIsViz highlights $\{a, b, c\}$ at frequency = 20% in the overview.

P15. When we select some prefixes of an itemset X in the detailed view, we can see from the overview how the frequency changes when truncating some items from X to form the prefixes of X .

P16. Similarly, when we select some extensions of an itemset X in the detailed view, we can see from the overview how the frequency changes when appending some items to X to form the extensions of X .

3. Extra interactive features of WiFIsViz

The above representation of WiFIsViz gives users insight into the general distribution of frequent itemsets; it also provides users with some relevant details. In this section, we describe some additional interactive features of WiFIsViz provided for user convenience.

Feature 1: query on cardinality. Recall that, by counting the number of nodes on each horizontal line in the detailed view, we can find all k -itemsets for a user-specified cardinality k . For user convenience and *interactive mining*, WiFIsViz allows the user to pose a query on cardinality, and it only shows k -itemsets. Figure 3 shows six buttons on the top for users to find itemsets of cardinalities 1, 2, 3, 4, 5 and/or 6.

Feature 2: query on frequency. With WiFIsViz, users can easily visualize all *frequent items* and/or *frequent itemsets*. Among them, users may interest in those of a certain frequency (or a certain range of frequencies). Hence, WiFIsViz provides users with an extra feature so that they can freely move the slider (on the left of the display area) up and down along the y -axis to show only those itemsets of a certain range of frequencies. By doing so, users can focus of those itemsets of their interest and avoid distraction from other uninteresting itemsets.

Feature 3: query on itemsets. WiFIsViz also allows users to select certain items of interest (e.g., promotional items in a store) and to pose queries on itemsets. Examples of these queries include (i) “find all itemsets containing *some* of selected items”, (ii) “find all itemsets containing *all* of the selected items”, and (iii) “find all itemsets *not* containing any of the selected items”.

Feature 4: details-on-demand. Details-on-demand [15] consists of techniques that provide more details whenever the user requests them. The key idea is that WiFIsViz gives users an overview of the entire dataset and then allows them to request details about selected parts by hovering the mouse over items. Specifically, WiFIsViz supports details-on-demand in the following ways: (i) When **the mouse hovers on a segment of a horizontal line** connecting two nodes (say, representing items b and d), WiFIsViz shows a list of itemsets containing both b and d . Selecting an itemset in the list instantly highlights the specific segment it is contained in, as well as both of its connecting nodes, so that users can see where the segment starts and ends. (ii) When the **mouse hovers over a node**, WiFIsViz shows a list of all itemsets contained in all the line segments starting or ending at this node. Selecting an itemset from the list instantly highlights the line it is contained in. (iii) When the **mouse hovers over a pixel** in the display, a small box appears showing the frequency and itemsets encoded by the mouse position. This is particularly useful when users need to see among the vast array of line segments what a particular point in the display refers to.

Feature 5: ordering of domain items. By default, WiFIsViz arranges the domain items (in the x -direction) in non-ascending frequency order. However, WiFIsViz also provides users with an option to arrange items in other orders. Having such an option is useful for *constrained mining*, in which users may want to arrange the items according to some constraints (e.g., put items of interest—say, promotional items—on the left and other items on the right of the x -axis).

Feature 6: visualization of other patterns. In addition to visualize frequent itemsets, WiFIsViz also provides users with options to visualize other patterns—such as **closed itemsets** and **maximal itemsets**. An itemset X is *closed*

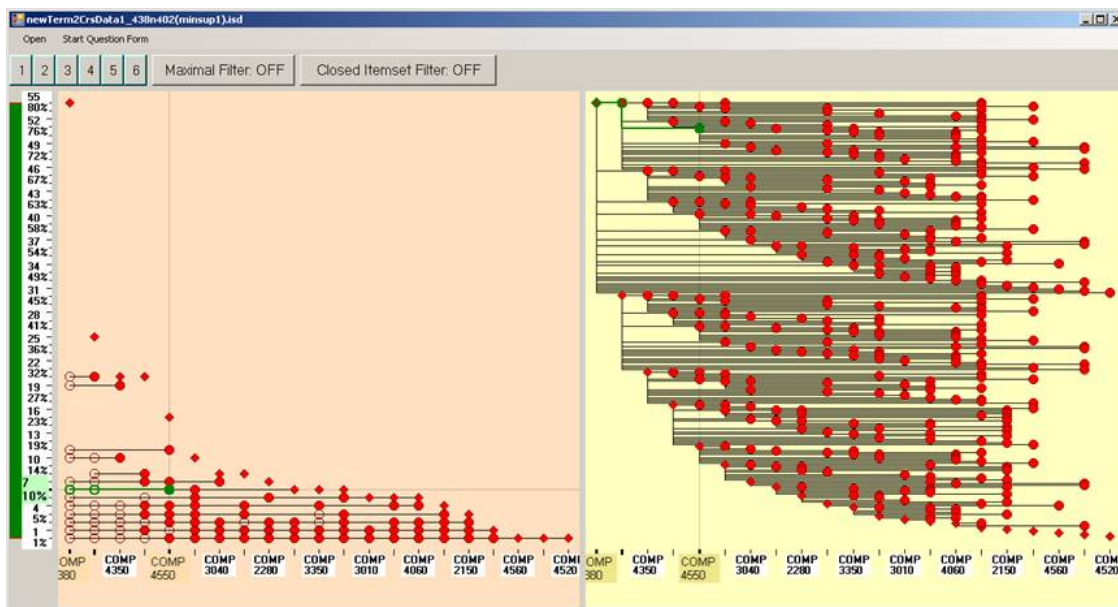


Figure 3. A snapshot of our proposed WiFISViz.

if there does not exist any proper superset of X that have the same frequency as X . By using this option (clicking the “Closed Itemsets” button), WiFISViz reduces the number of displayed itemsets without losing any information.

Similarly, an itemset Y is *maximal* if there does not exist any proper superset of Y that is also frequent. With this option, WiFISViz further reduces the number of displayed itemsets because the collection of maximal itemsets is always a subset of the collection of closed itemsets, which in turn is always a subset of the collection of frequent itemsets.

4. Evaluation results

In this section, we show our evaluation results of the proposed WiFISViz. We first compare the functionality of our WiFISViz with an existing system. We then conduct a user evaluation on the effectiveness of visualization provided by WiFISViz.

4.1. Functionality

Here, we compared our proposed WiFISViz with an existing visualizer, FISViz [11], by considering many different real-life scenarios. For each scenario, we determined whether these systems can handle the scenarios. If so, we examined how these systems display the mining results. Our results showed that WiFISViz was effective in all these scenarios, including the following two scenarios.

Scenario 1. What kinds of vegetables are in the biggest shopper basket? Here, we aim to find itemsets having the highest cardinality. To answer this question using FISViz, we look for polylines with the highest number of nodes. However, polylines may be bent and cross each other. As shown in Figure 1, users may *not* be able to easily identify which line segment belongs to a polyline (and which does

not). In contrast, to answer this question using WiFISViz, we can easily find horizontal lines with the highest number of nodes from the detailed view.

Scenario 2. What kinds of fruits are purchased together by $y\%$ of customers? In FISViz, itemsets are represented by polylines, which may be bent or cross each other. Hence, users may *not* be able to visualize all itemsets of frequency $y\%$, especially in a very dense or clustered area of the display. In contrast, it is not so crowded in WiFISViz due to projection compression. Users can easily identify and select from the overview the horizontal line of frequency $y\%$. As the two views of WiFISViz are linked, the corresponding lines (representing all itemsets of frequency $y\%$) are highlighted. This gives the kinds of fruits that are purchased together by $y\%$ of customers.

4.2. Readability and interaction

To assess the effectiveness of conveying frequent itemset relationships, we carried out a *user evaluation* with WiFISViz. The evaluation was primarily case-based, within which several types of users were required to solve many different questions based on the visualization of a given dataset (e.g., a dataset consisting of courses taken by upper-level undergraduate students in the current academic term). Figure 3 shows a snapshot of WiFISViz displaying frequent itemsets mined from such a dataset. Participants were required to identify a set of relationships and make decisions based on their observations from the visualization. We recruited about 20 participants and separated them into two groups: (i) those who have data mining background and (ii) those who do not. However, none of the participants (regardless which of the two groups) was exposed to any form of visualization for frequent itemsets—including WiFISViz.

To test the expressivity of our visualization, we formulated two sets of questions. The first set of questions was aimed at assessing whether participants could simply pick out useful information from viewing the data. This would inform us whether the itemsets were easily visible, whether the frequencies of items were easily identifiable, and whether the layout of the two views made it easy for participants to tease out information. Sample questions included “Which two courses are most frequently taken?” and “Which three courses are taken together by four students?”. Our evaluation results showed that all the test participants were able to identify the basic meaning behind the representations (e.g., frequency was assigned to the y -axis in the overview, courses was assigned to the x -axis). The majority of participants were able to correctly answer most—and some even correctly answer all—of the questions. Interestingly, while we gave the participants a multiple-choice list to help them identify the answers, none of the participants used the multiple-choice to guide their selection. Instead, their curiosity in understanding the visualizations led them to answer the questions by looking at the two views of WiFIsViz first, and then confirming their answer with one of the choices provided. Moreover, the evaluation results were similar in both groups (with vs. without data mining background). This showed the ease of readability with WiFIsViz. The results also suggested that, with very little training, participants felt comfortable to use the visualizations. Furthermore, they were able to quickly assimilate the representations to the point at which they were able to answer all questions adequately.

The second set of questions required participants to perform some level of more analytical reasoning with the visualization. These included open-ended questions such as “If you were to schedule exams of these courses, which pair of courses would you avoid scheduling on the same day?”. Participants essentially provided answers that suggested they could extract a significant amount of the visualization and make sense of it. Our results showed that the participants were able to make the best use of visualization for answering these questions. For example, a participant reported that “it would make sense to schedule exams for disjointed courses (e.g., COMP 4520 & 4560). Examination scheduling is quite a complicated optimization problem, but somewhat easily read from the current visualization provided by WiFIsViz”.

Moreover, the participants also found that, while most of the interactive features were not essential (e.g., one could count the number of nodes to determine the cardinality of an itemset), they were very helpful and convenient to users.

To summarize, participants were able to easily capture the relationships in our datasets, to identify answers to most questions with very high accuracy, and to perform analytical reasoning with the visualization. The results were highly

promising in that they confirmed our intuitions about the effectiveness of our proposed visualizer—WiFIsViz.

5. Conclusions

Most of the frequent itemset mining algorithms return the mined results in textual forms, which can be very long and difficult to comprehend. As “a picture is worth a thousand words”, it is desirable to have visual systems. However, many existing visual systems were not designed to show frequent itemsets. To improve this situation, we proposed and developed a powerful frequent itemset visualizer called *WiFIsViz*, which provides users with frequent itemset information that is easy to understand. Specifically, *WiFIsViz* gives a quick intuitive overview of all the itemsets and their frequencies. It also provides in-depth details of interesting itemsets. Our evaluation results showed that *WiFIsViz* was effective in visualizing frequent itemsets, which in turn helps users in answering a board range of questions for real-life applications and assists them in making appropriate business or management decisions.

Acknowledgement. This project is partially supported by NSERC (Canada) in the form of research grants. Also thanks all those who participated in the user evaluation.

References

- [1] R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In *Proc. VLDB 1994*, pp. 487–499.
- [2] M. Ankerst et al. Visual classification: an interactive approach to decision tree construction. In *Proc. KDD 1999*, pp. 392–396.
- [3] P. Appan, H. Sundaram, and B.L. Tseng. Summarization and visualization of communication patterns in a large-scale social network. In *Proc. PAKDD 2006*, pp. 371–379.
- [4] S. Berchtold, H.V. Jagadish, and K.A. Ross. Independence diagrams: a technique for visual data mining. In *Proc. KDD 1998*, pp. 139–143.
- [5] J. Blanchard, F. Guillet, and H. Briand. Interactive visual exploration of association rules with rule-focusing methodology. *KAIS* **13**(1), pp. 43–75, 2007.
- [6] C. Brunk, J. Kelly, and R. Kohavi. MineSet: an integrated system for data mining. In *Proc. KDD 1997*, pp. 135–138.
- [7] D.A. Keim. Information visualization and visual data mining. *IEEE TVCG* **8**(1), pp. 1–8, 2002.
- [8] E. Keogh et al. Intelligent Icons: integrating lite-weight data mining and visualization into GUI operating systems. In *ICDM 2006*, pp. 912–916.
- [9] Y. Koren and D. Harel. A two-way visualization method for clustered data. In *Proc. KDD 2003*, pp. 589–594.
- [10] L.V.S. Lakshmanan, C.K.-S. Leung, and R.T. Ng. Efficient dynamic mining of constrained frequent sets. *ACM TODS* **28**(4), pp. 337–389, 2003.
- [11] C.K.-S. Leung, P.P. Irani, and C.L. Carmichael. FIsViz: a frequent itemset visualizer. In *Proc. PAKDD 2008*, pp. 644–652.
- [12] C.K.-S. Leung and Q.I. Khan. DSTree: a tree structure for the mining of frequent sets from data streams. In *Proc. ICDM 2006*, pp. 928–932.
- [13] C.K.-S. Leung, Q.I. Khan, and T. Hoque. CanTree: a tree structure for efficient incremental mining of frequent patterns. In *Proc. ICDM 2005*, pp. 274–281.
- [14] C.K.-S. Leung, M.A.F. Mateo, and D.A. Brajczuk. A tree-based approach for frequent pattern mining from uncertain data. In *Proc. PAKDD 2008*, pp. 653–661.
- [15] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *Proc. VL 1996*, pp. 336–343.
- [16] C. Stolte, D. Tang, and P. Hanrahan. Query, analysis, and visualization of hierarchically structured data using Polaris. In *Proc. KDD 2002*, pp. 112–122.
- [17] L. Yang. Pruning and visualizing generalized association rules in parallel coordinates. *IEEE TKDE* **17**(1), pp. 60–70, 2005.
- [18] K. Zhao et al. A visual data mining framework for convenient identification of useful knowledge. In *Proc. ICDM 2005*, pp. 530–537.