ScienFIST.org

INTERNATIONAL JOURNAL OF INFORMATION TECHNOLOGIES, ENGINEERING AND MANAGEMENT SCIENCE

Data Visualization in 2D/3D with Usage of Force-driven Algorithms

Martin Mazúch, Jozef Kostolný, Alexander Brezáni
Faculty of Management Science and Informatics, University of Zilina, Slovak Republic
martin.mazuch@fri.uniza.sk, jozef.kostolny@fri.uniza.sk, alexander.brezani@fri.uniza.sk

Abstract

Force-driven algorithms offer a straightforward and intuitive approach to visualizing data structures and social networks in two dimensions. Their simplicity facilitates rapid data presentation and fosters understanding of inter-element relationships. However, limitations such as visual clutter, overlapping elements, and a lack of realism hinder their effectiveness in capturing intricate data nuances. To address these challenges, this paper proposes a novel three-dimensional visualization method. By combining traditional force-driven techniques with spherical displays, data normalization, clustering, spring coupling, and gravitational interactions, we aim to create a robust framework for representing complex data structures while preserving hierarchical and structural information. This approach promises to enhance data visualization capabilities and provide deeper insights into the underlying relationships within networks.

Keywords: Force-driven algorithms, data visualization, 2D/3D environment, Game engines

Introduction

In this introduction to the problem of force-driven algorithms in 2D space, we focus on their fundamental strengths, such as simplicity and comprehensibility, which make them an ideal tool for visualizing different types of data structures and social networks. These algorithms allow for a fast and intuitive presentation of data, which is important for an effective understanding of the relationships between elements.

On the other hand, the two-dimensional space also carries with it some limitations. A cluttered graph can lead to overlapping vertices and edges, which negatively affects its readability. Another drawback is the lack of realism, which can limit the ability to capture fine details in the relationships between vertices.

This introduction leads to the idea that visualization in three-dimensional space offers a wider scope for capturing complex relationships but requires new techniques and optimizations. Therefore, in this chapter we focus on an innovative approach that combines traditional methods with new techniques such as spherical displays, data normalization, clustering, spring coupling, and gravitational interactions. This approach not only allows for a better representation of the data, but also considers the hierarchical and structural properties of data networks, thus providing a robust tool for visualization in three-dimensional space.

Materials and methods

Visual analysis in two-dimensional space

The advantages of force-driven algorithms in 2D space lie in their simplicity and comprehensibility. This form of visualization provides fast and intuitive data presentations, which is advantageous for various applications of social network visualizations or data structures. However, despite these advantages, 2D space also carries disadvantages. The limited space can lead to overlapping vertices and edges, which reduces the readability of the graph. Also, the lack of realism can limit the ability to capture details of relationships and distances between vertices.

The forces between nodes can be calculated based on their theoretical distances in the graph, determined by the lengths of the shortest paths between them. For example, the algorithm of Kamada and Kawai [1] uses spring forces proportional to the theoretical distances of the graph. In general, force-oriented methods define an objective function that maps each graph layout to a number v, which represents the energy of the layout. This function is defined such that low energies correspond to arrangements in which neighboring nodes are close together based on some pre-specified distance and in which non-adjacent nodes are well spaced [2].

Advantages of visualisation in 2D space

The traditional way of visualizing graphs and networks, including big, connected data, has several advantages. However, based on the nature of the algorithms and the limitations of the display area, these advantages are gradually lost as the complexity of the relationships or the size of the network increases. The major positives of 2D visualization include:

- Simplicity and clarity: in 2D space, force-driven graphs are often simple and easy to understand. This is especially the case for graphs up to the size of tens of vertices.
- Speed of computation: algorithms for 2D space tend to be faster compared to their 3D equivalents. There are also algorithms that can be easily parallelized, which further increases the speed of creating the resulting visualization.
- Wide Use: In 2D space, force-driven algorithms are suitable for a wide range of applications.

Disadvantages of visualization in 2D space

In analyzing visualizations in 2D space, we focused mainly on identifying shortcomings of two-dimensional visualization that can be addressed by adding an additional dimension and implementing a force-based approach using the existing capabilities of physical environments. The main shortcomings of 2D visualization include:

- Limited space: vertices and edges can overlap in 2D space, which can degrade the readability of the graph. This shortcoming is crucial when there are clusters whose diameter is small.
- Lack of realism: 2D visualization may not fully capture the relationships and distances between vertices. This is especially the case when the length of edges must be considered as a real value expressing, for example, the spatial length.
- Limited interaction possibilities: in the case of closely spaced vertices, the need for user interaction may not be obvious, e.g. performing a scene rotation operation to get better information about the interrelationships in the graph.

Results and discussion

Methods and algorithms for 2D visualization

Basic force-based approaches have a fundamental limitation imposed by the number of vertices to be rendered/processed. Their use is limited to small graphs and the results are poor for graphs with more than a few hundred vertices. There are several reasons why traditional power-oriented algorithms do not work well for large graphs.

One of the main obstacles to their scalability is the fact that the physical model usually has many local minima. Even with the help of sophisticated mechanisms for avoiding local minima, basic power-oriented algorithms are not able to consistently produce good distributions for large graphs. Similarly, barycentric methods do not work well for large graphs, mainly due to resolution issues (for large graphs, the minimum vertex separation is very small, leading to unreadable visualizations) [2].

In the late 1990s, several techniques emerged extending the functionality of power-directed methods to graphs with tens of thousands and even hundreds of thousands of vertices. The main idea behind these approaches is the multilevel decomposition technique, where the graph is represented by a series of progressively simplifying structures and arranged in order from the simplest to the most complex. These structures can be coarser graphs (as in the approach of Hadany and Harel [3], Harel and Koren [4] and Walshaw [5]) or vertex filterings as in the approach of Gajer, Goodrich and Kobourov [6].

Visual analysis in three-dimensional space

Data visualization in three-dimensional space plays a key role in the analysis and understanding of complex relationships in various research areas and applications. New challenges in data visualization require continuous improvement of existing algorithms and the creation of new approaches to ensure efficient and intuitive perception of three-dimensional data structures. In this chapter, we present an innovative algorithm based on the principle of force-driven algorithms that enables visualization in 3D space, but also integrates spherical representations, normalization, clustering, spring coupling, and gravitational interactions to create a robust tool for visualizing complex data networks.

The algorithms are widely used in the field of data visualization and are based on the principle of simulating physical forces between vertices of a graph. Their effectiveness in two-dimensional spaces is well documented, but transferring these approaches to a three-dimensional environment requires new techniques and optimizations. In our approach, we have enhanced traditional force-driven algorithms with spherical representations, allowing for a better representation of the relationships between vertices in a three-dimensional space.

In addition, we implemented normalization techniques that allow efficient manipulation of the range of values and improve the consistency of the visual presentation. The use of clustering allows us to identify and visualize subsets of data with similar characteristics, while spring couplings and gravitational interactions add additional dynamics to the visualization, providing a better interpretation of the data graph structure.

In the following sections of the chapter, we describe in more detail the proposed algorithms, their implementation, and the results that are obtained when visualizing heterogeneous data networks in a three-dimensional space. This approach provides a new perspective on data visualization, considering not only geometric aspects but also hierarchical and structural properties of the data.

Force-driven algorithms in 3D space bring the advantage of a more accurate representation of distances between vertices. This capability enables the creation of visualizations that are more readable and structured, which is significant for detailed explorations of data patterns. With advantages in 3D space come disadvantages. Algorithms in 3D space are more computationally complex, which can cause a greater computational burden. The increased complexity of visualization can in turn lead to loss of detail and increased difficulty of interactions with the environment.

Benefits in 3D Space:

- Accurate distance: in 3D space, force-driven algorithms have more space to accurately represent distances between vertices.
- Better readability.

Disadvantages in 3D Space:

- More computationally complex: Algorithms for 3D space are often more computationally intensive.
- Increased complexity of the visualization: By increasing the space, the complexity of the visualization can increase.
- Greater Number of Interactions: Users may need a greater number of interactions to accurately view and manipulate the three-dimensional space.

Classical force-driven algorithms are limited to computing the graph layout in Euclidean geometry, typically, some more recent implementations in for larger values of . However, there are cases where Euclidean geometry may not be the best choice, which is also the result of the findings of the dissertation. Certain graphs have a structure that is better visualized using a different geometry, such as on the surface of a sphere or on an anuloid. In particular, representations of 3D mesh can be displayed on a sphere in order to map textures or to impose them on a torus without crossing. In addition, certain non-Euclidean geometries, namely hyperbolic geometry, have been found to have properties that are particularly suited for the arrangement and visualization of large classes of graphs [7] [8]. Based on the above findings, Kobourov and Wampler describe extensions of power-directed algorithms to Riemann spaces [9].

Preprocessing the data and the network as a whole is an essential part of visualization, and preprocessing can be performed with different focuses. In most cases, the data needs to be preprocessed for the sake of cleaning up erroneous values, normalization or other operations that will subsequently make the data easier to work with as described by the authors [10]. A second possible reason for performing data preprocessing may be to try to lighten the algorithms used to compute information from the data, crawl it, or lighten the hardware resources for further processing.

Network preprocessing

Nowadays, there are many algorithms that are designed to set a comprehensible and clear network shape using algorithms to compute basic network characteristic metrics. The resulting computed values are then used as a starting point for algorithms to lay out the network in space. In 2D space, this approach is quite common and easy to apply to different kinds of data, but in the 3D case, calculating metrics alone is not enough for us. Metrics determining network properties in 3D space can be misleading, so it is necessary to properly select metrics that the algorithms can work with efficiently in 3D. In an early version of our work, we use force-driven algorithms for which the underlying metrics are just a way to speed up the computation of the final position of the network.

Data preparation

Visualizing data in three-dimensional space plays a key role in analyzing and understanding complex relationships in various research areas and applications. New challenges in data visualization require continuous refinement of existing algorithms and the creation of new approaches to ensure efficient and intuitive perception of three-dimensional data structures. In this chapter, we present an innovative algorithm based on the principle of force-driven algorithms that enables visualization in 3D space, but also integrates spherical representations, normalization, clustering, spring coupling, and gravitational interactions to create a robust tool for visualizing complex data networks.

Scaling edge values

Scaling edge values is the process of adjusting the values that represent the relationships between vertices in a graph. The goal of this process is to achieve a uniform range of values that is suitable for visualizing and analyzing graph data. Scaling edge values can be important because it allows for better comparison and interpretation of relationships between vertices without unfairly plotting values due to their large variations. Scaling edge values is essential for achieving consistent and interpretable results when visualizing in 3D.

Normalization of parameters for edge values

Parameter normalization for edge values is a process that aims to achieve a uniform and standardized range of parameter values that represent the relationships between vertices in a graph. This procedure is often used when preparing data for graph visualization, whether in 2D or 3D space. Standardization is important for several reasons:

- Simplifying comparisons: normalized values allow for easier comparison of relationships between vertices because they are scaled to a standard interval.
- Minimizing the impact of outliers: Normalization helps to eliminate the impact of outliers that might otherwise unfairly affect the visual display of the graph.
- Ensuring consistency: Normalisation helps to achieve a consistent range of values for different edges of the graph, making them easier to interpret.

Conclusion

In conclusion, the analysis of force-driven algorithms in both two-dimensional (2D) and three-dimensional (3D) spaces reveals a balance between simplicity and complexity in data visualization. While 2D visualization offers significant advantages such as simplicity, clarity, and computational efficiency, it is limited by its inability to accurately represent complex relationships and distances, particularly as the size and complexity of the data increase. This limitation often leads to issues such as overlapping vertices and edges, which can diminish the readability and effectiveness of the visualization.

On the other hand, transitioning to 3D visualization addresses many of the shortcomings of 2D space by providing a more accurate representation of relationships and distances between vertices. This allows for a richer and more detailed exploration of data, especially in complex networks. However, the advantages of 3D visualization come with their own set of challenges, including increased computational complexity and the potential for more intricate interactions required to manipulate and interpret the visualized data.

The evolution of force-driven algorithms from 2D to 3D space reflects the need for continuous adaptation and innovation in data visualization techniques. By integrating advanced approaches such as spherical representations, normalization, clustering, and spring coupling, these algorithms have been enhanced to provide more robust tools for visualizing complex data networks in three-dimensional space. This development underscores the importance of not only preserving the geometric and structural integrity of the data but also ensuring that the visualization remains interpretable and useful for in-depth analysis.

Overall, while 2D visualizations remain valuable for their simplicity and ease of use, 3D visualizations offer greater potential for detailed analysis, particularly in contexts where spatial relationships and data complexity are critical. The ongoing refinement of algorithms to better handle the demands of 3D visualization will continue to play a key role in advancing the field of data analysis and visualization.

References

[1] T. Kamada a S. Kawai, „An algorithm for drawing general undirected graphs," Inform. Process. Lett., zv. 31, pp. 7-15, 1989.

[2] S. G. Kobourov, „Force-Directed Drawing Algorithms," rev. Handbook of Graph Drawing and Visualization, R. Tamassia, Ed., CRC Press, 2013, pp. 383-408.

[3] R. Hadany a D. Harel, „A multi-scale algorithm for drawing graphs nicely," Discrete Applied Mathematics, zv. 113, %1. vyd.1, pp. 3-21, 2001.

[4] D. Harel a Y. Koren, „A fast multi-scale method for drawing large graphs," rev.

Graph Drawing. GD 2000. Lecture Notes in Computer Science, zv. 1984, J. Marks, Ed., Berlin, Heidelberg, Springer, 2002, pp. 179-200.

[5] C. Walshaw, „A multilevel algorithm for force-directed graph drawing," Journal of Graph Algorithms and Applications, zv. 7, %1. vyd.3, pp. 253-285, 2003.

[6] P. Gajer, M. T. Goodrich a S. G. Kobourov, „A fast multi-dimensional algorithm for drawing large graphs," Computational Geometry: Theory and Applications, zv. 29, %1. vyd.1, pp. 3-18, 2004.

[7] J. Lamping, R. Rao a P. Pirolli, „A focus+context technique based on hyperbolic geometry for visualizing large hierarchies," rev. Proceedings of Computer Human Interaction, 1995.

[8] T. Munzner, „H3: Laying out large directed graphs in 3D hyperbolic space," rev. Proceedings of IEEE Symposium on Information Visualization, 1997.

[9] S. G. Kobourov a K. Wampler, „Non-Euclidean spring embedders," IEEE Transactions on Visualization and Computer Graphics, zv. 11, %1. vyd.6, pp. 757-767, 2005.

[10] S. M. Anouncia, H. A. Gohel a V. Subbiah, Data Visualization: Trends and Challenges Toward Multidisciplinary Perception, Springe Singapore, 2020.