Master research Internship

Bibliographic report

Study of the interest of 3D immersive visualization in decision-taking

Domain: Human-Computer Interaction, Social and Information Networks

Author:
Rémi Sahl

Supervisor:
Laurent Brisson
Thierry Duval
Lab-STICC

Abstract: This paper will examine the potential and the difficulties behind the development of an immersive graph visualization for the study of social networks. At first, we will have a look into social network analysis methods and the visualization process to understand how visualization helps domain experts in their search. We will then propose an overview of classical graph visualization and data analysis techniques, and of some existing work in graph visualization. Finally, we will investigate how and why immersive graph visualization would be more effective than common 2D displays. This state of the art will allow to understand the stakes and the necessity of researching new tools and techniques for a better use of immersive VR technologies in the graph visualization process. This work precedes an internship that will consist in developing an immersive graph visualization prototype with the Unity3D platform.

Contents

1 Introduction **1**

2 Social networks and visualization **2**
- 2.1 Social network analysis . . . 2
- 2.2 The visualization process . . . 3

3 Graph visualization **4**
- 3.1 Graph layout techniques . . . 5
- 3.2 Graph clustering techniques . . . 6
- 3.3 Exploration and manipulation . . . 8
- 3.4 Tools and libraries . . . 10

4 Immersive virtual reality for graph visualization **10**
- 4.1 Virtual reality and visualization . . . 11
- 4.2 Layout and interaction methods for immersive graph visualization . . . 12

5 Conclusion **15**

1 Introduction

In a world where the amount of data never ceases to expand, it becomes harder and harder to properly harness and understand this huge quantity of information. This "information overload" [10] leads to various problems related to the issue of finding and interacting with the relevant information in a relevant way. It is difficult for us to perceive and understand this massive amount of data without finding the right manner of representing it. This is where visualization and visual analytics techniques become useful as they help presenting the data in a human-understandable form, with which we can interact, and from which we can infer more information [27]. Visualization allows to distinguish patterns and investigate the data more easily. This paper will focus on graph visualization, which is a type of representation widely used throughout history to help model data elements and their internal connections. The science of graph visualization and analysis has already been extensively researched as it can have "*many applications, such as social networks, Internet communications, paper citations, and biochemical pathways*" [2].

Among the many applications of graph visualization, our research will focus on social network (SN) analysis. SN analysis probably originated with the works of Auguste Comte in the beginning of the XIVth century [3] and has been particularly expanding lately with the apparition of massive online networks such as Facebook, LinkedIn, etc... Several visualization softwares have been designed to visualize and analyze these kinds of graph. For example, we will see that softwares such as Gephi or Graphviz and can be used to help detect communities, predict the diffusion or rumors, or identify roles in SN.[1]

However, these visualization techniques are mainly intended to traditional 2D representations and as immersive Virtual Reality (VR) becomes more and more accessible, immersive visualization shows more and more potential. Would the use of modern immersive technologies such as Head-Mounted Displays really benefit to the overall performance of the visualization [28]? If so, in which ways? One should be aware that, although it might enhance our perception of complex shapes [24], the addition of a third dimension is not always helpful to visualization [15] [28]. Recent work such as [12] & [14] show that the immersive visualization of such graphs would probably be more effective that traditional 2D displays. However, graph visualization techniques in immersive VR still need research and development, as we need new and more effective methods and tools to examine and interact with the data in an immersive context.

This paper will examine the potential and the difficulties behind the development of an immersive graph visualization for social networks analysis. At first, we will have a look into social network analysis methods and the visualization process to understand how visualization helps domain experts in their search. We will then propose an overview of classical graph visualization and data analysis techniques, and of some existing work in graph visualization. Finally, we will investigate how and why immersive graph visualization would be more effective than common 2D displays. This state of the art will help understand the stakes and the necessity of researching new tools and techniques for a better use of immersive VR technologies in the graph visualization process. This work precedes an internship that will consist in developing an immersive graph visualization prototype with the Unity platform.

2 Social networks and visualization

In the last decades, social media has taken a more and more important part of our lives. Its' expansion has risen new questions on how we can analyze the huge amounts of data harvested from social networks. Although social network analysis can be applied to much smaller systems than those of social media giants such as Facebook or Twitter, these huge collections of data have given new challenges to domain experts. Visualization tools are very useful to help overcome these challenges and help apprehend this data in a more understandable form. This section will begin by presenting an overview of a few social analysis concepts which will help us grasp the necessity of visualization. We will then study the different aspects of this visualization process.

2.1 Social network analysis

> "*Social media mining is the process of representing, analyzing, and extracting meaningful patterns from data in social media, resulting from social interactions.*"[29]

For this process to be completed, visualization is essential. It is usually displayed in the form of a graph where the nodes are actors and the edges interpersonal connections. As mentioned in [1], SN graphs are essentially containing three variables:

- the structural variable that describes the connections between actors (edges)
- the composition variable that contains information about each actors individually (nodes)
- the affiliation variable that indicates extra information about the actors (for example groups to which they belong to)

A lot of knowledge can be inferred from this representation, especially with the appropriate tools and algorithms to guide the visualization process. These few concepts of SN analysis will help us understand what use of visualization is made by experts to infer new information from the data gathered:

Community detection An application of SN analysis is the detection of communities. These techniques are used to detect an aggregation of actors that are linked in a particular way (which is not always as obvious as in the case of Facebook groups). For example, this process was used in 2010 to discover a network of 30 millions infected computers transmitting botnets all over the web [29]. "*The problem that community detection attempts to solve is the identification of groups of vertices that are more densely connected to each other than to the rest of the network*", as mentioned in [19]. This identification relies on clustering techniques that will be developed in section 3.2. As pointed out in [1], "*Several methods have been developed to find clusters in a graph, or which is equivalent, to find communities in a social network.*"

Information diffusion The study of the diffusion of information within a social network is also helped with appropriate visualization tools. For these phenomenon to be modeled and studied, edges transmit information from one end to an other (in a directed or undirected way depending on the type of the graph), and nodes are defined with states that can change upon time. The notion of temporality is also central to these studies and needs to be taken into account in the visualization.

Guha et al.[6] used these properties to study the propagation of trust and distrust, Kempe et al.[11] tried to maximize the spread of influence through a social network, and Nekovee et al.[16] studied the theory of rumour spreading.

These few concepts show some potential applications in the study of social network analysis and highlight the necessity of visualization tools. The next subsection will present the unfolding of the visualization process and will help us understand how visualization is helpful in the study of such concepts.

2.2 The visualization process

In order to study the relevant technical aspects behind visualization, we need to understand how domain experts make use of such a software and how it is helpful to their study. It is important to pinpoint the interdisciplinarity inherent to the field of visualization, as it shows that visualization should be adapted to the task at hand and requires the use of domain experts in its development.

A definition of graph visualization [2]:

> "*Graph visualization is a sub-field of information visualization. It usually refers to representation of interconnected nodes arranged in space and navigation through a visual representation to help users understand the global or local original data structures. It is a complex field, since it draws on ideas from several intellectual domains: computer science, psychology, semiotics, graphic design, cartography, and art. It also makes the task of analyzing a set of data with relations full of challenges.*"

As mentioned in [26], visualization highly depends on the data which needs to be represented. Contingent upon the number of independent or dependent variables, and on the type of each of these variables (discrete, continuous, vectors, booleans, and so on...), the data can be represented in various ways (Charts, graphs, pies, etc..). The visualization process is usually established in four stages. These four stages consist in:

- Collecting data
- Preprocessing the data into a representable form
- Rendering a visualization
- Perceiving the data through human-computer interaction and exploration, and in so inferring new knowledge

The following illustrations presents a way of mapping the visualization process.

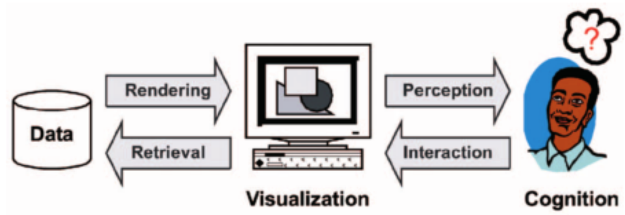


Figure 1: The visualization process, as presented in [25]

The perception of the rendered data is critical and one should keep in mind the cognitive aspects behind human computer interaction and our mental representation of data. L.Grammel et al.[4] asked novices to instinctively do the work of a visualization expert by exploring a fictitious set of data and reporting on their insight. This work highlighted our natural way to represent and interact with data as well as some common interpretation problems we can face when visualizing it. This issue is also highlighted in the work of Ware et al.[27] who tries to understand and model the way we perceive different types of data. This is also a major concern of Tory et al.'s[25] paper who studied the impact of colors, shapes, and visual effects on our perception. These aspects are primordial in the development of any visualization and these choices should not be made without reflection. Keim et al.[10] also mentioned the importance of human cognition in their work on visual analytics.

High visual complexity, due to a high number of data items, occlusion, and very spiky line-chart profiles

Unfamiliar visualization types, e.g. scatter plots

Inappropriate scaling of measurement mappings (axes, color, size) and **inappropriate width/height ratio**

Inappropriate size of the visualization

Difficulties understanding semantics of measurements, including the selection operation (e.g. average, sum)

Inappropriate levels of abstraction, either too high or too low

Readability problems, e.g. bright colors, small font sizes and inappropriate positioning of labels and legends

Missing numbers

Figure 2: Information visualization novices' interpretation problems [4]

In social network analysis, the data is represented as graphs that sometimes can be very complex. Therefore, domain experts need ways of visualizing such complex networks while avoiding the common interpretation problems above. These tools need to facilitate the perception of the data and allow interaction with it in an intuitive way. The next section will describe how graph visualization techniques are applied to help guide this visualization process. This description will help identify efficient techniques commonly used on traditional displays (flat-screens), and in so, to find clues on how could these techniques be applied in immersive VR.

3 Graph visualization

This section will present an overview of graph visualization techniques meant to improve the performance of the visualization process. The main objective behind those techniques is to facilitate the perception of the user on the data set and its exploration. We will begin by looking at graph layout techniques which allow the drawing of the graph in the most easily understandable way. We will then see that various clustering mechanisms are used to help us manage the visualization of large graphs such as those of some social networks. Finally, we will take a brief look at existing visualization softwares, as they present good examples of implementation of these techniques.

3.1 Graph layout techniques

This subsection presents graph layout techniques. Once the data has been collected (first stage of the visualization cycle), we need to represent it into a graph. In order to do so, we first need to define the elementary items of a graph (nodes and edges), before computing spatial positions of these elements. As mentioned in section 2, the cognitive aspects are to be taken into account in this stage as those spatial positions need to be optimized for better readability and comprehension.

Graph definition [13]

> "*A graph refers to a set of vertices (nodes) and a set of edges (i.e., links) that connect pairs of vertices. [...] Furthermore, attributes can be attached to vertices and edges, e.g., to denote their type, size, or some other application related information.*"

When managing large graphs, multiple issues are to be dealt with in order for the user to properly perceive the representation of the data. The drawing needs to be clear and precise or else it can easily become confusing. These general aesthetic rules highlight what should be prioritized when placing the nodes and edges, and were described in [2] and [1]:

- Distribute evenly the nodes and edges in space [2] [1]
- Avoid edge crossing [2] [1]
- Minimize the bend along edges [2] [1]
- Maximize symmetry [2] [1]
- Keep edges parallel to coordinate axes [1]
- Maximize the angle between two consecutive edges incident to a node [1]

Illustration 3 highlights how easier it becomes to read the same graph once it has been properly laid out. This illustration is also a good example of the aesthetic rules above. Note that this graph is relatively small compared to some sociograms (social network graphs). The notions of size (number of nodes) and density (number of edges relative to the maximum potential number of edges) are used in graph visualization to quantify graph dimensions.[13]

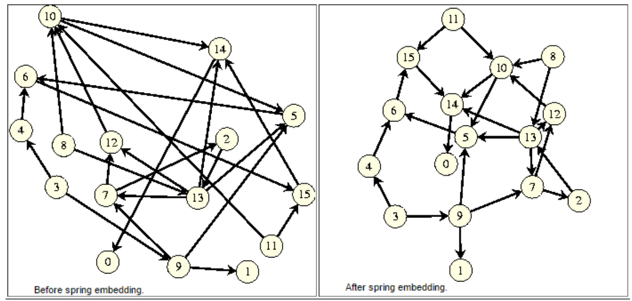


Figure 3: Before and after applying a layout algorithm[1]

In order to display the graph in the most effective manner according to the aesthetic rules defined above, a large collection of algorithms are used and specifically designed depending on the application and type of graphs. It is quite difficult to establish which of these algorithms should be used and in which case, as they all present different manners of organizing the data. Some researchers have tried to establish comparisons of graph readability [13] using eye-tracking techniques for example. Usually, classic algorithms are extended and modified to fit the specifics of a visualization task, or of certain types of graphs (graphs can be defined as cyclic, acyclic, direct or undirected for example). Therefore, a great deal of implementation of graph layout algorithms can be found and an extensive survey would be well beyond the scope of this work. These are probably the most commonly used algorithms for the laying out of graphs:

- Tree layouts [22] [23] [13] [2] : These approaches are mainly used for directed graphs and are based on Sugiyama's approach [22]. Nodes are assigned to layers and positioned accordingly.
- Force-directed layouts [2] : These are physics-based approaches in which repulsive forces are assigned to nodes and attraction forces are assigned between endpoints of edges. These forces tend to stabilize the system into a state that fits the aesthetic rules above. These algorithms usually give good results but do not scale well with size. See figure 3 for an example of such a layout.
- Constraint-based layouts [13] : These approaches are extensions of force-directed layouts with the addition of constraints on alignment or other aspects.

However, even with the best layout algorithms, the visualization of very large and complex graphs can easily become too cluttered and difficult to apprehend. The next subsection will present data analysis and clustering techniques that are used in graph visualization to help addressing these issues.

3.2 Graph clustering techniques

Besides the drawing of the graph, the complex structure of the data needs to be taken into account. In SN analysis, nodes represent people and edges represent interpersonal connections but these nodes and edges may contain information relative to the personal information of the actors

[1] http://www.leda-tutorial.org/en/discussion/ch05s03s08.html - 12/2016

and to the type of connection between them. A way to account for this additional data in the visualization is to represent additional dimensions with non-euclidean variables such as color or shape [8]. This issue of not being able to apprehend high-dimensionality systems is called the curse of dimensionality. To overcome this issue, one needs to transform the original data set into one that is defined upon less dimensions using dimension reduction techniques [2].This process can be facilitated by implementing various data analysis techniques such as clustering mechanisms. These mechanisms are particularly effective in graph visualization as they address the minimization of the size of the data to be viewed. They can be used over several aspects of the graph, semantic or structural, depending on the needs. [23]

For example, in *survey on graph visualization*, Cui et al. presented node-clustering techniques used to reduce visual clutter [2]. Different levels of clustering can be implemented over the structural aspect which can allow users to see various levels of detail depending on this factor. This aspect is essential in the navigation aspects of visualization tools when dealing with large amounts of data and will be additionally detailed in section 3.3.

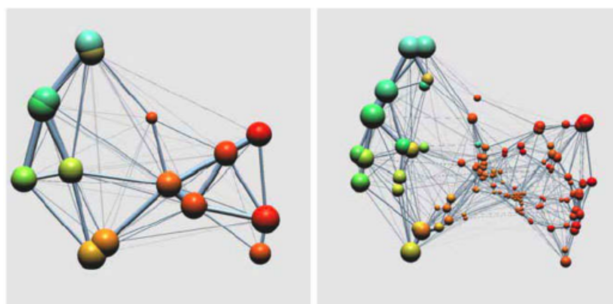


Figure 4: Levels of clustering, as presented in [2] (Node clustering)

Just as for layout algorithms, a great variety of clustering algorithms are used in graph visualization. These techniques usually base themselves on the notion of distance, which can be defined as euclidean or else, to compute a proximity value between nodes and then assign them to clusters. We will now have a quick overview of the most common graph clustering algorithms:

- Graph partition methods [1]: These methods divide the graph into a predefined number k of groups. These groups are defined by trying to minimize the number of edges between them.

- Hierarchical clustering [1]: "*These techniques place each object in a single group and gradually merge those atomic groups into larger and larger clusters until all the objects are in a single cluster.*" Hierarchical clustering is particularly useful for revealing hierarchical structures if they exist.

- Partitional clustering [1]: These approaches are similar to graph partition methods techniques as they assign nodes to a predefined number of k groups except the main criteria is the distance from each node to the group's center of mass. The most famous of these algorithms is the k-means algorithm in which nodes are initially randomly assigned, and then iteratively reassigned upon this criteria until the cluster configuration remains stable.

- Spectral methods [1] [2] [20]: These methods base themselves on algebraic and arithmetic techniques of matrices analysis to compute pair-wise dissimilarity or distance matrices. These methods can not be applied to any type of graphs and generally need to be completed with other clustering techniques such as k-means.

- Divisive algorithms [1]: Divisive algorithms are based upon the notion of centrality. Edges are removed iteratively based on this criteria. The goal is to find the most central edges which connect clusters of nodes between them and thus discover the underlying structure of the graph.

- Grid-based : Hinnerburg et al. proposed a grid-based form of clustering that is said to be particularly efficient when dealing with high-dimensional data [8]. These techniques divide the spatial environment into a grid. All data points which fall into the same grid cell are then aggregated and treated as one object.

As mentioned before, these various techniques of clustering can be used over several aspects of the visualization, hence the need for personalized visualization depending on the task at hand. The visualization process should allow full control over these aspects as they are no better clustering technique than an other. Such algorithms should be implemented and adapted only to guide the user in his search for the right representation.

3.3 Exploration and manipulation

Besides the aspects relative to the laying out and the clustering of the graph, appropriate user interaction is essential to any visualization software. In order for the user to properly understand the data, he needs to be able to manipulate and explore the visualization. This subsection will focus on the essential aspects of user interaction in graph visualization. What should be implemented to enhance the user experience while visualizing the data? How do these techniques help with the overall performance of the visualization?

Once the graph is drawn, the user needs to be able to evolve in the scene and explore it. In this exploratory phase, the graph is usually static and only the user moves around, inside, and outside the graph to perceive it from different points of view. Navigation needs to be easy and intuitive and the user should be able to do so without difficulties. This allows him to understand the complex structure of the data to be visualized, both locally and globally.

Layout algorithms usually become insufficient to handle great amount of data and become easily unreadable (as shown in figure 5 below). As pointed out in [23], zooming mechanisms need to be applied to understand the entire graph as they allow more precise inspection of its' semantic value. However, "*It is obvious that when we zoom in a graph we lost information about the context and when we zoom out we lost details. The best known solution is to provide a overview of the entire graph, which can guarantee a minimal context*" [23]. Zooming mechanisms can be improved by various clustering mechanisms and levels of clustering, as shown in figure 4, can be associated to them. Note that these mechanisms are very efficient for better readability according to the aesthetic rules above.

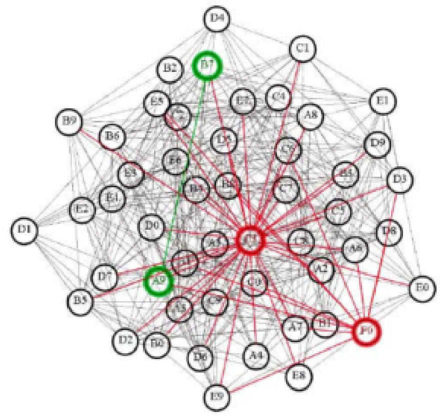


Figure 5: A visualization of a graph containing 50 vertices and 400 edges [2]

The user should also be able to interact over several other aspects of the visualization in an easy and intuitive way. These aspects should of course be defined upon what should be visualized and how. For example, other forms of user interaction could include layout algorithm modifications, clustering algorithm modifications, data selection, manipulation of the graph position and rotation, search bars, menus, etc... Shneiderman et al. [21] highlighted the importance of design in good human-computer interaction and presented guidelines towards appropriate design and good user experience which should be applied in visualization. From a software-engineering point of view, classic rules of conception such as design-patterns should be applied for the conception of robust and extensible software.

The variety of graph layout and data analysis algorithms presented in the beginning of section 3 offers a great deal of possibilities in personalized visualization. However, the way we should represent the data and interact with it, often requires content-based knowledge. As mentioned in [23], the visualization and interaction components should be designed with active participation of domain experts of each field of application. "*The environment would be adapted by means of personalized toolbars and plug-ins, depending on the type of graph that the user wants to interact with*"[23]. The work of [20] introduces several important interaction techniques when manipulating 2D graphs:

- **Simulated lenses** allow detail on demand
- **Graph selection** is usually made through pointer selection or query selection
- **Structural & topological navigation** allow users to show and hide part of the graphs to obtain better visibility
- **Degree of interest navigation** computes and highlights interesting nodes based on a subset of the data and on various criterion
- **Multivariate graphs** such as SN graphs contain rich metadata from which we can define types and organize the graph accordingly.

3.4 Tools and libraries

This subsection will present a few of the most famous solutions to visualize and manipulate graphs. Such visualization solutions usually take most of the problems studied above into account and are obviously interesting to mention in our study. As pointed out in [1], there exists a huge variety of visualization tools and libraries. Figure 6 presents a few of the most popular works and their capabilities:

Type	Name	Language	Capabilities
Library	GraphViz [Ellson et al., 2001]	C++	This is a graph visualization software that implements some general graph layout algorithms. It is not designed to be used in an interactive environment.
	OGDF [Chimiani et al., 2007]	C++	This library contains several graph drawing and manipulation algorithms, providing a graph manipulation engine.
	Prefuse [Heer et al., 2005]	Java	This is an extensible framework for manipulating graphs in interactive ways. The framework includes connections to Web search engines and SQL storage interfaces.
Software	Cytoscape [Shannon et al., 2003]	Java	This is a software project created for analyzing protein interaction and molecular networks. However, it can be used with complex networks. It is extensible through plugins.
	Gephi [Bastian et al., 2009]	Java	This is a graph manipulation tool created to provide users with a simple but powerful tool. It includes several layout and partitioning algorithms and graph measures as well. It is extensible via plugins.
	Tulip [Auber, 2003]	C++	This is also a graph manipulation tool created to manipulate any kind of network. It provides a set of general manipulation algorithms, including layout, partitioning and coloring. It is extensible via plugins.

Figure 6: Tools and libraries for manipulating graphs and complex networks, as presented in [1]

In their paper, Jacomy et al.[9] presented forceAtlas2, a force-directed dynamic layout algorithm designed for the Gephi software. This algorithm is essentially similar to classical force-directed layouts presented in 3.2, except it supports additional constraints relative to the dynamic aspects and the Gephi interface. "*Its implementation of adaptive local and global speeds gives good performances for network of fewer than 100000 nodes, while keeping it a continuous layout (no phases, no auto-stop), fitting to Gephi user experience.*" This approach is interesting as it shows that the layout method of a visualization should be conceived while keeping in mind how the user would interact with it.

However, most graph visualization tools and techniques described in this section were designed to be implemented on traditional 2D displays. With the recent advances in immersive VR technologies, the next section will show that 3D immersive graph visualization shows great promises of performance, even though redesigning these tools and techniques to fit this new form of display is going to be a difficult task.

4 Immersive virtual reality for graph visualization

If visualization performance is directly related to what the user can perceive in the visualization process, would immersive VR help improve this performance and offer new ways of interacting with the data? In which ways? If so, how would one extend the traditional graph visualization methods to an immersive context? These questions become more and more relevant as VR technologies becomes more and more advanced and accessible. However, much work still remains to be done as we have not yet found the perfect ways of visualizing and interacting with graphs in this context. In this section, we will begin by examining the potential of immersive VR in visualization. In a second part, we will present a few 3D layout and interaction methods for immersive graph visualization.

4.1 Virtual reality and visualization

VR systems have been developed for a few decades and many devices where conceived to immerse users in virtual reality. This technology can take many forms such as caves, domes, and 3D stereoscopic devices, but was not always an affordable solution until recently. These last few years, VR headset (Head-mounted displays) technologies have been widely democratized and this technology will most probably continue to be greatly improved in the incoming years. There is a lot of progress to be made both in hardware and software and the possibilities are far from being fully exploited. Immersive VR is very efficient in term of immersion and presence within the scene and this could be very helpful to the visualization process.

The early work of Ware et al.[28] in 1996 has demonstrated that immersive virtual reality is very helpful for visualizing information nets in 3D and this confirms our intuition that immersive graph visualization would be more effective in an appropriate immersive VR environment. His work was already conclusive back then and this is very promising in light of the recent improvements of the technology. Tory et al.[24] evaluated *visualization task performance on 2D, 3D, and combination displays* and highlighted the fact that the addition of a 3rd dimension with appropriate cues (shadows for example) are very effective for approximate navigation and relative positioning. However, 2D displays are generally better for precise navigation and positioning which makes the use of combination displays relevant. The visualization community is beginning to realize the potential in VR visualization and recent works such as the one of Kwon et al. [12] shows that immersive visualization is proving more effective than those on classical 2D displays. However, more research is clearly needed as we need more effective ways of displaying and interacting with the data. Many solutions are yet to be explored for us to optimize visualization tasks. "*We hope our work and findings will encourage others to join this exciting area of study so they can help accelerate the development of usable technologies to meet the growing demand of more effective tools for examining and analyzing large complex data.*[12]" More research is still clearly needed to furthermore examine how to interact with such immersive environments [7] as we are not yet accommodated to designing and implementing immersive VR programs.

However, the use of immersive VR technologies such as Head Mounted-Displays comes with a few issues that needs to be considered when implementing such systems. The work of McIntire [15] on "*the (possible) utility of stereoscopic displays for information: the good, the bad, the ugly*" shows that our brain requires immediate feedback on immersive interaction and that delays between motion and scene update for example can induce simulator sickness. Stereoscopic 3D displays are known to sometimes cause discomfort, eyestrain, fatigue effects, and nausea on the user. The exact causes of these effects are not yet fully understood but are mostly due to perceptual or cognitive conflict. Results showed that self-motion parallax (via headtracking) [7] greatly improves overall user's perception and comfort and is a good step towards better immersive technologies. These issues restrict the possibilities of motion within the environment in an immersive context.

4.2 Layout and interaction methods for immersive graph visualization

Because the expansion of VR technology is still recent, we need to find new and appropriate ways to layout the graph, navigate through it, and interact with it. Although various approaches exist, we are not yet sure which technique is more efficient. These approaches are usually extensions of classical graph layout algorithms, clustering algorithms, and interaction techniques for 3D immersive visualization. More work is clearly needed to fully exploit the potential of VR technologies in graph visualization, and we need to find new manners of implementing the aesthetic and user interaction guidelines for graph visualization (mentioned in section 3) in this new context. This subsection will present a few existing work of immersive graph visualization while trying to identify relevant techniques of layouts and interactions.

In their work, *Immersive Dynamic Visualization of Interactions in a Social Network*, Greffard et al. [5] presented an immersive graph visualization using a dynamic force-directed layout similar in some aspects to forceAtlas2. Their solution (see figure 7 below) allowed an immersive visualization from an external point of view and was conclusively tested on a social network analysis application. "*It allowed to understand, through the identification of articulation points linking different communities, the formation and merging of communities.*"

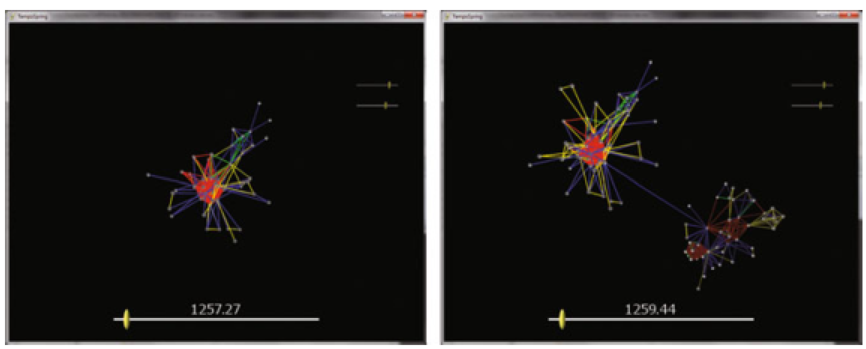


Figure 7: An immersive visualization of graph using a force-directed layout algorithm [5]

More recently, Kwon et al. [18] [12] proposed a new immersive 3D layout solution using De Boor's algorithm to project classical 2D graph layouts onto the surface of a sphere. This projection makes all nodes equally visible to the user (see figure 8 below). This layout technique removes the need for navigation around the graph since the user is motionless at the center of the sphere.

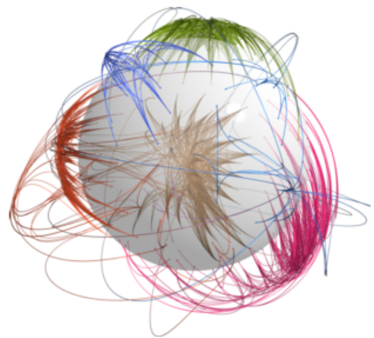


Figure 8: Spherical projection of a graph [18]

In another recent work, Liluashvili et al. introduced interactome-CAVE (iCAVE), *an open source tool for immersive 3D visualization of complex biomolecular interaction networks*. Although this work is intended for the visualization of biomolecular networks, it is a very complete system and a very inspiring work of immersive graph visualization. It implements multiple layout algorithms, clustering mechanisms, user-interface functionalities, and many exploration and manipulation features, to allow the user to chose the most suited techniques for his task. For the layout of the graph, i-Cave implemented an extension of classical force-directed layouts to 3D (figure 9), a semantic levels layout that uses a combination of force-directed and layered displays to compute 3D positions (figure 10), and a hemispherical layout which is kind of similar to Kwon et al.[18]'s approach (figure 11).

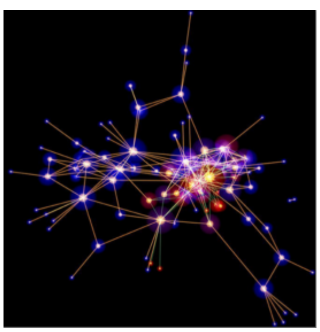


Figure 9: iCAVE's force directed layout [14]

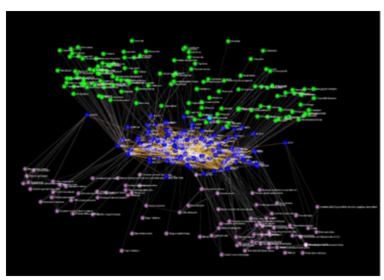


Figure 10: iCAVE's semantic levels layout [14]

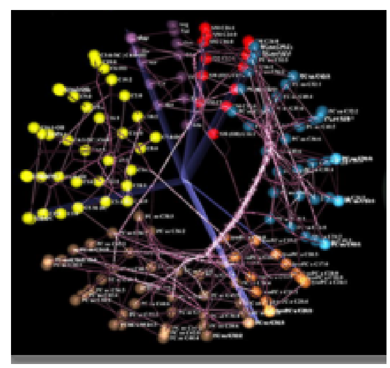
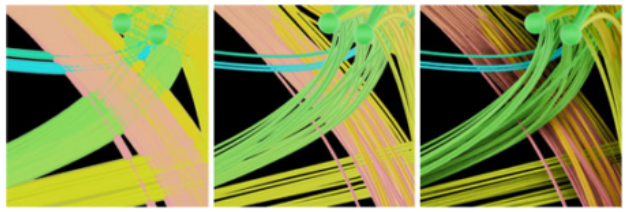


Figure 11: iCAVE's hemispherical layout [14]

Exploration In immersive VR, exploration needs to take a particular form, as motion can easily cause motion-sickness. Self-motion parallax with head-tracking, along with graph manipulation techniques, might be sufficient for the exploration stage of the visualization process (with the appropriate layout). Self-motion parallax is particularly effective to perceive complex 3D shapes and helps locating oneself in space [7]. In their solution (figure 12), Kwon et al. chose not to allow the user to move around the graph, as nodes are evenly distributed around the user. Their efficient use of depth routing and illumination techniques (see figure 12 below) added an impression of depth on the spherically-projected graph. This practical technique allowed them to highlight (while displaying them in the foreground) selected or interesting nodes or edges (which pairs well with the graph interaction guidelines presented in 3.3).



(a) without depth routing (b) with depth routing (c) with depth routing and illumination

Figure 12: Depth routing and rendering techniques

Interaction Interaction techniques also need to be re-evaluated, as they are mainly intended for 2D representations. For example, we need to find other ways of selecting nodes or edges in a graph than classical pointer selection if we use hand-tracked devices. This is a weak spot in immersive VR and new techniques are still being researched. An example of such research is the work of Nguyen et al.[17], who developed a new *direct manipulation techniques for immersive 3D environments* that facilitates the manipulation of objects in an immersive VR system. Such technique tries to account for the fact that "*human beings have difficulties in performing accurate manipulation tasks or in keeping the hand motionless in a particular position without the help of external devices or haptic feedback.*" For these reasons, Kwon et al.[12] & Greffard et al.[5] preferred to use traditional keyboard and mouse controls in their immersive visualization. This choice however, leaves promising potential of immersive interaction unexploited (with hand-tracked controls). Liluashvili et al. [14] tried to implement both traditional and hand-held controls in their visualization, as shown in figure 13 below, and allowed the user to zoom or rotate the network for exploration.

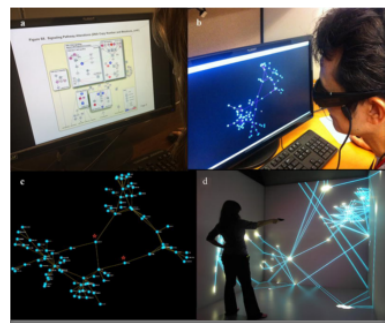


Figure 13: the iCave visualization [14]

These various approaches are only the beginning of a promising use of immersive VR technologies in graph visualization. This section showed that such form of display could really benefit to the performance of the visualization. However, immersive graph visualization could make much better use of the potential of immersive VR and we need to develop new techniques of layout and interaction adapted to this form of display.

5 Conclusion

Social network analysis is a complex field and domain experts make efficient use of visualization softwares to help them understand and represent their networks. The efficiency of this visualization process relies on the way the user perceives the rendered representation of the data. Immersive VR technologies are particularly efficient in the perception of complex 3D shapes, and in that sense, their use could really facilitate this visualization process. Therefore, new tools of graph visualization adapted to this form of display could really benefit the analysis of large and complex networks, such as those of social networks. This potential use of immersive VR technologies in graph visualization is beginning to be researched as this technology is becoming more and more advanced and affordable.

However, the development of such immersive graph visualization tools comes with new challenges and constraints associated to this new form of display. The traditional methods of laying out the graph were meant for flat screen display and we need to develop new algorithms that allow an optimized and clear visualization. Clustering techniques are also relevant in immersive graph visualization and useful for this process. We also need new exploration and manipulation methods that fit this new context. Although it is possible to enjoy the benefits of stereoscopic 3D displays with traditional keyboard and mouse controls, a better use of hand-tracked devices could potentially be made, especially as this technology is also being developed. We described a few attempts made in immersive graph visualization but we believe that there is still improvements to be made and that new techniques of interaction and layout need to be more extensively researched.

This state of the art would have allowed us to understand the interest of investigating new ways to implement such an immersive graph visualization software that could be used in the study of graphs such as those of social networks. It helped identify important features one would need to implement in the development of such a visualization. This knowledge will help in the work of the upcoming internship, during which we will attempt to do so using the Unity3D platform and immersive VR technologies.

References

[1] Juan David Cruz-Gomez. *Socio-semantic networks algorithm for a point of view based visualization of on-line communities.* PhD thesis, 2012.

[2] Weiwei Cui and Huamin Qu. A Survey on Graph Visualization. *The Journal of infectious diseases*, 208:NP, 2013.

[3] Linton C Freeman. *The development of social network analysis.* Number January 2004. 2004.

[4] L Grammel, M Tory, and M Storey. How Information Visualization Novices Construct Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952, nov 2010.

[5] Nicolas Greffard, Fabien Picarougne, and Pascale Kuntz. Immersive Dynamic Visualization of Interactions in a Social Network. pages 255–262. 2012.

[6] R Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 403–412, 2004.

[7] Chris Hand. A Survey of 3D Interaction Techniques. *Computer Graphics Forum*, 16(5):269–281, 1997.

[8] Alexander Hinneburg and Daniel a Keim. Optimal Grid-Clustering: Towards Breaking the Curse of Dimensionality in High-Dimensional Clustering. *International Conference on Very Large Databases (VLDB)*, pages 506–517, 1999.

[9] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS ONE*, 9(6):1–12, 2014.

[10] Daniel Keim, Gennady Andrienko, Jean-daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. Visual Analytics: Definition, Process, and Challenges. In *Information Visualization*, pages 154–175. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[11] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03*, page 137, 2003.

[12] Oh Hyun Kwon, Chris Muelder, Kyungwon Lee, and Kwan Liu Ma. A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1802–1815, 2016.

[13] T Von Landesberger, a Kuijper, T. Schreck, J. Kohlhammer, J.J. van Wijk, J.-D. Fekete, and D.W. Fellner. Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. *Computer graphics ...*, 30(6):1719–1749, 2011.

[14] Vaja Liluashvili, Selim Kalayci, Eugene Flouder, Manda Wilson, Aaron Gabow, and Zeynep H. Gümüş. iCAVE: an open source tool for immersive 3D visualization of complex biomolecular interaction networks. Technical report, jun 2016.

[15] John P McIntire and Kristen K Liggett. The (possible) utility of stereoscopic 3D displays for information visualization: The good, the bad, and the ugly. In *2014 IEEE VIS International Workshop on 3DVis (3DVis)*, pages 1–9, 2014.

[16] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. Theory of rumour spreading in complex social networks. *Physica A: Statistical Mechanics and its Applications*, 374(1):457–470, 2007.

[17] Thi Thuong Huyen Nguyen, Thierry Duval, and Charles Pontonnier. A New Direct Manipulation Technique for Immersive 3D Virtual Environments. *ICAT-EGVE 2014 : the 24th International Conference on Artificial Reality and Telexistence and the 19th Eurographics Symposium on Virtual Environments*, page 8, 2014.

[18] Oh-Hyun Kwon, Chris Muelder, Kyungwon Lee, and Kwan-Liu Ma. Spherical layout and rendering methods for immersive graph visualization. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, volume d, pages 63–67. IEEE, apr 2015.

[19] Symeon Papadopoulos, Yiannis Kompatsiaris, Athena Vakali, and Ploutarchos Spyridonos. Community detection in social media performance and application considerations. *Data Mining and Knowledge Discovery*, 24(3):515–554, 2012.

[20] Robert Pienta, James Abello, Minsuk Kahng, and Duen Horng Chau. Scalable graph exploration and visualization: Sensemaking challenges and opportunities. *2015 International Conference on Big Data and Smart Computing, BIGCOMP 2015*, pages 271–278, 2015.

[21] Ben Shneiderman and Catherine Plaisant. Designing the User Interface: Strategies for Effective Human-Computer Interaction, 2010.

[22] Kozo Sugiyama, Shojiro Tagawa, and Mitsuhiko Toda. Methods for Visual Understanding of Hierarchical System Structures. *IEEE Transactions on Systems, Man and Cybernetics*, 11(2):109–125, 1981.

[23] JB Tolosa and JEL Gayo. Visualization, Navigation and Edition of Graph Structured Semantic Information. *Josebarranquero.Com*, 2008.

[24] Melanie Tory, A.E. Kirkpatrick, M.S. Atkins, and T. Moller. Visualization task performance with 2D, 3D, and combination displays. *IEEE Transactions on Visualization and Computer Graphics*, 12(1):2–13, jan 2006.

[25] Melanie Tory and T. Moller. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics*, 10(1):72–84, jan 2004.

[26] Melanie Tory and Torsten Möller. Rethinking visualization: A high-level taxonomy. *Proceedings - IEEE Symposium on Information Visualization, INFO VIS*, pages 151–158, 2004.

[27] Colin Ware. *Foundation for a Science of Data Visualization.* 2012.

[28] Colin Ware, David Hui, and Glenn Franck. Evaluating stereo and motion cues for visualizing information nets in three dimensions. *ACM Transactions on Graphics (TOG)*, 15(2):121–140, 1996.

[29] Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. Social Media Mining An Introduction. *Cambridge university Press*, page 382, 2014.