Abstract—The video game market is a complex system where the relationships between game titles and publishers is convoluted. While some publishers are responsible for a large number of titles others seem to be less successful. In an attempt to understand the characteristics and evolution of the video game industry, we analyze the market using social network analysis. For this purpose, we build a heterogeneous network of the video game industry at different levels. Moreover, to interpret the behavior of the generations of game consoles, we build these networks for each console, i.e., PS3, Xbox 360 and Wii, and analyze the three generations they went through. We perform a comparative temporal analysis of the evolution of the industry by investigating 14 video game consoles as five generations. Finally, we try to provide insights and future directions regarding the video game market.

Keywords-Complex Networks; Game Consoles; Social Networks; Video Game Industry

I. INTRODUCTION

Analyzing how the industry has changed over the time can help predict where the industry is going in the future. The video games industry is projected to bring in revenues over 60 billion dollars in 2011 [7]. This value has been increasing since video games gained mass acceptance when arcades were popping up in the late 70’s and early 80’s. An industry this big has many people wanting to gain entry. However, without any knowledge of how the industry is shaped it can be a pretty daunting task.

It has been shown that visual representations may provide a better idea of the characteristics of a system than just raw numbers [2]. In this paper, after building corresponding network, we visually map out the video game industry.

Console video games systems are gaming systems that players typically use at home. All video games are released for a specific console. For example, a game that runs on the Playstaton 3 will not run on an Xbox 360. Games, however, can have separate releases that allow the same title to play on both consoles. For instance, Madden 2012 has a release for both the Xbox 360 and the Playstation 3. Games are created by a developer and released by a publisher. A publisher may release games for several developers, and developers may release several titles. The relationship between titles, developers, publishers and consoles can provide an understanding of the games industry.

The video games industry has been around for over 30 years. As technologies change, the industries behind them may change as well. It is unknown how the relationship between developers, publishers and consoles has changed over the course of this relatively young industry. In this paper, we compare network topologies of different generations of video game console industry to understand how the industry has changed over the time.

The first generation of video games from the 1970’s and 1980’s will be called classic video games. The first console in this category is the Atari 2600, which was released in 1977. Atari 2600 has been credited as popularizing cartridge based home gaming. The second console included in this group is the Nintendo Entertainment System (NES), which was released in 1985. At one point, the NES was the best selling console of all time. The final console in this category is the Sega Master System, which was also released in 1985 as a direct competitor to the NES.

The current generation of video games consoles include products from Nintendo, Microsoft, and Sony. All of these consoles were released within the last 6 years. The first console included in this category is the Nintendo Wii, which was released in 2006. This console revolutionized game play with its motion capturing controller and is the best selling console of all time with over 87 million systems sold world wide. The second console in this category is the Sony Playstation 3 (PS3), which was released in 2006. The final console included in the current generation is Microsoft’s Xbox 360, which was released in 2005 as the successor to Microsoft’s first consumer product, the Xbox.

In addition to the classic and current video games, we will also analyze other generations in between. The following are the five generations of 14 video game consoles:

1) Atari 2600, Nintendo Entertainment System, and Sega Master System
2) Super Nintendo, Sega Genesis, and Turbo Grafx 16
3) Nintendo 64 and Playstation
4) Playstation 2, Xbox and Gamecube
5) Nintendo Wii, Playstation 3, and Xbox 360

The rest of the paper is organized as follows. Section II covers background information and previous work to justify the methods for displaying and analyzing the video game industry. Section III explains the methods of data collection and data analysis. Section IV reveals the results, and Section V provides an overview of the results. Finally, the paper is concluded in section VI.
II. RELATED WORK

Earlier research has shown that both graphical and numerical analysis are valid ways of comparing social networks.

A. Graphical Analysis

It has been shown that it is easier to recognize differences when they are represented by position, size or color as opposed to a strictly numerical display [1]. For instance, the global seed industry network between 1996 and 2008 has been animated to reveal the change in the industry over the years [2]. The network showed the parent companies as nodes with size proportional to the number of smaller companies they owned. Over time, acquisitions and newcomers into the industry are visualized. Major characteristics of the industry can be obtained from watching the short video.

Looking at soft drink selection at a local convenience store may give the appearance that there are many choices, however a 2008 study [3] shows that most of these choices trickle back to one of a few major parent companies. In particular, 89% of the choices belong to one of three parent companies. Although this statistic seems to be fairly significant, when represented as a visual network where the node sizes are based on the amount of soft drink brands they own, major players become very obvious. We expect the overall network topology of the soft drink industry to be very similar to the topology of the video game industry over the past 25 years.

B. Numerical Analysis

Although a visual comparison between complex networks is useful, it is important to analyze network measures. There are several well known metrics which are widely utilized in social network analysis. Size is one of the most basic properties of a network and quantified by the number of nodes \( n \) and the number of edges \( e \). The basic characteristic to infer graph connectivity is average node degree \( \bar{k} = 2n/e \). The degree \( k \) of a node is the number of edges that are adjacent to the node. In-degree represents the frequency edges that point to a node while out-degree represents the frequency edges originating from the node.

Clustering coefficient is the measure of how well the adjacent (neighbors) of a node are connected. The neighbor set \( ns \) of a node \( a \) is the set of nodes that are connected to \( a \). If every node in the \( ns \) is connected to each other, then the \( ns \) of \( a \) is complete and will have a clustering coefficient of 1. If no nodes in the \( ns \) of \( a \) are connected, then the clustering coefficient of \( a \) will be 0. High clustering coefficient is the indicator of small-world effect along with small average shortest path.

Out-degree and clustering coefficient will be used when comparing two generations of video games. In-degree is not useful in this study as it will be the same for all levels due to the hierarchical nature of the network. Shortest paths are also irrelevant due to the nature of the topology. Using these measures in conjunction with the visual representation may give a good indication of the type of change that has occurred in the video game industry over the past 25 years.

III. METHODOLOGY

In order to analyze the video game industry over the past 30 years, several items needed to be completed. First, the data needed to be collected. Secondly, the data needed to be analyzed in a graphical manner, and finally the data needed to be analyzed in a numerical manner. The procedures for achieving these goals is outlined in the following subsections. The overall goal of this analysis is to predict certain characteristics of the video game industry for the next generation of consoles. If these predictions prove to be accurate, they could help a newcomer to the video game industry figure out a business plan that will be in line with the next generation.

A. Data Collection

The most important piece of this project was to get the representative data about game consoles. Although the data does exist, it is scattered across the Internet, and not available as a complete dataset. We gathered the data from two websites. www.games-db.com website was founded in 2003 and is dedicated to archiving data relating to video games, including console, publisher, developer and titles - all of which are needed for this project. Although this site is comprehensive for the data is has, it is not updated frequently enough to contain the latest releases. This website was used to collect data on earlier generation of consoles, except for the current generation. The data for the current generation of consoles was pulled from www.ign.com website. IGN is a global leader in three business verticals: media, digital distribution, and game technology. Their data for past consoles is not as complete as their information on the current generation.

Neither of these websites offer direct access to their data. As a result, a custom web scraper was written from scratch to grab this data. The utility, Video Game Grabber (VG2), was developed in C#. The HttpWebRequest API was used to get the root webpage for each console, and then custom routines were written to get the information for all of the titles for that particular console. If we were not able to determine a particular game’s publisher or developer, the fields for that game were filled with the string ERROR. Examples of errors were when custom items such as footnotes were added to the websites to further describe some aspect of the game. When these were encountered, the information was initially listed as ERROR, and a modification was performed on the data by hand. The output of the VG2 were comma separated files (csv) as it is easy to import that into common programs for analysis. The data for each console was saved in a separate file and contained the game title, the publisher and the developer.

B. Graphical Analysis

As discussed earlier, a graphical representation of a complex network has shown to be beneficial [1]. For this study, all of the consoles’ networks were generated based on the relationships of console, publisher, developer and title. Two different tools were used to generate visual representations of the networks.
1) **Pajek**: is a program for analyzing networks and is freely available at the project website [5]. Unfortunately, Pajek does not take data in a csv format. As a result, the VG² software was modified to translate the csv file into a .net file that Pajek could interpret. The VG² program traversed the entire csv, and assigned a unique identifier to each node in the graph. Separate lists were maintained for each of the publishers, developers, and titles, while a global identifier was used to uniquely distinguish them from each other when generating the .net file. Once all the unique nodes were identified, the Vertices section of the .net file were written out. After that, the relationships between all the nodes were identified, and the Edges section of the Pajek .net file were written out. The weights were all set to a 1. Using degree partitions, the network was drawn using the Kamada-Kavai Energy layout.

Although Pajek colored the nodes based on their degree, it was a little difficult to determine the important nodes. If the size of the nodes was altered such that nodes with high degree are drawn larger, it would be easier to identify who the important players are.

2) **GUESS**: is the Graph Exploration System [6]. As GUESS uses a different file format than Pajek, the data needed to be converted. The VG² software was modified to convert a csv to .gdf. The file format is similar, however instead of using numbers that uniquely identify the node, the .gdf format uses strings. In order to not confuse GUESS, strings in the network were stripped of their special characters. The nodes were then drawn with sizes proportional to the out degree of each node. Similarly, physics layout was used to draw the topology.

C. **Numerical Analysis**

The graphical analysis will provide a nice overview of how the industry has changed over the years, but it will not directly show any results based on the statistics of the network. In order to quantitatively look at the change, a number of metrics must be looked at. The specific metrics are listed throughout this section. The metrics will be averaged for 3 consoles for each generation, then plotted to see if there can be any type of line or curve fit to the data. If the results show that there can be some kind of curve fit to the data, it will be important because it may, potentially, not only show where the industry has been, but where it might go in the future. This information could be used by someone who is starting up a game development company and wants to anticipate what the structure will be like in the next generation of consoles.

One tool used for the numerical analysis of the change of the video game industry will be Pajek. Pajek provides several network measures to analyze a network. Additionally, when needed other tools will be written from scratch.

As shown in [4], degree distribution is a crucial metric to analyze a network. The network of the video games is a directed network with publishers linking to developers who link to titles. Because of this structure, out degree will be used to analyze the video game network. Just as in the analysis of the clustering coefficient, the values for each console of a particular generation will be averaged to generate a line of best fit. Because there are two different pieces of information that are relevant in this network, we will analyze the out degree at 2 different levels of the hierarchy. First we will look at the out degree at the publisher level to reveal the number of developers that typically work for each publisher. Then we will analyze the out degree at the developer level to reveal the number of titles created by each developer.

We will also analyze the placement of data. The degree distribution will show the average values at each level, but we will need to look a little bit deeper to see how these averages were achieved. The out degree of the developers and the out degree of the publishers will be plotted for all members of each group. These will be compared from the first generation to the current generation to see if either generation follows power law.

IV. **RESULTS**

In this section, we provide analysis of the collected data using the above mechanisms.

A. **Graphical Results**

Although 14 consoles were analyzed for this research, 2 graphical representations are shown here. They are the network map for the Nintendo Entertainment System (Figure 1), and the Nintendo Wii (Figure 2). The nodes of the graphs are sized according to their out degree. Larger nodes are connected to more nodes. The largest 10 nodes in the network are colored red, as well as labeled. From these graphs it is interesting to see the overall increase of the number of nodes. The current generation of consoles have so many nodes, that it becomes difficult to convey the underlying structure of the networks. The overall structure of the networks remains unchanged, the console industry is still a hierarchal structure.

![Fig. 1. NES Network](image)
B. Numerical Results

Figure 3 shows the degree distribution at both the Publisher and the Developer level of the hierarchical network for each console analyzed. The data is organized in sequence of console release with the Atari 2600 on the far left and the Wii on the far right. The spike in the middle section is due to a large number of Titles/Developer for the original Playstation console. The original Playstation checked in with an average of 6.01 titles per developer.

Figure 4 shows the degree distribution at both the Publisher and the Developer level of the hierarchical network averaged for each generation of consoles. This is a weighted average with each console for each generation given a weight equal to its share of the total number of nodes for that generation. For the third generation, only the original Playstation and the Nintendo 64 were analyzed. The Playstation accounted for 81.98% of the nodes in the network for that generation, and thus the spike remains when everything is averaged.

Figure 5 shows the out degree required to be in the top 10 nodes for that particular console. In order to make the graphical representations more readable, the top 10 nodes for that console were colored red. This chart breaks that graphical representation down a little further showing the numbers required to be in the top 10 for each console. Recent consoles are stabilizing around the 20 connection mark.

Figures 6 and 7 show the distribution of developers for the classic and current generation in a log log plot. The distributions resemble power law degree distributions where most developers generate few titles and few developers generate a large number of titles. Moreover, while the NES among the classic generation differs from others the consoles in the current generation show similar patterns. This indicates that the developer community around each console is similar in terms of produced titles.

Figures 8 and 9 show the distribution of publishers for the classic and current generation in a log log plot. These plots also exhibit power law patterns. Moreover, similar to the developer analysis, the classic generation has different distributions while the current generation exhibits the same distribution. This also indicates that the publisher community around each console is similar in terms of developers.
The graphical representation of the data shows how much the industry has grown over the past 30 years. Looking at the sheer number of nodes in the recent consoles shows how crowded the industry is. Simple graphs such as the one for the NES console easily show the hierarchical structure, however when looking at a recent consoles like the Nintendo Wii, the number of nodes appear as a large blob of data. If this trend continues, the next generation of consoles will probably be much more crowded.

The industry as a whole seems to be settling on 2 developers per publisher and 2 titles per developer per console on average. Early on, these numbers seemed to fluctuate, but as the industry matures, the volatility has subsided and as a result these values appear to be where the future consoles will go. A company wishing to enter the video game industry must prepare itself to only be focussed on 2 major projects over the lifetime of a console, either by writing two game titles if you are a developer, or by publishing for two developers if you are a publisher.

The increase in the number of nodes required to be in the top 10 nodes for a console can be interpreted with the consolidation of the industry. For current consoles, a value of 20 connections is required when the average is just 2. Joining this group will take considerable effort. Although stabilizing, this value still seems to have some growth, and to join ranks of the top 10 in the next generation, companies will need to have more than 20 connections.

The log log plots show a power law degree distribution for the current generation of consoles when looking at the developer and publisher levels of the network structure. The classic generation however is not as smooth power law as the current generations. Moreover, the degree distributions for the

V. DISCUSSION
current generation are very similar. If this type of distribution is to continue and the goal is to find a publisher, it may be a good idea to choose one of the few publishers who accept many developers. It can also be perceived that if an individual developer plans many titles, they all may not be completed unless the developer is one of the few that can produce that many titles.

VI. CONCLUSION AND FUTURE WORK

This paper presents a first attempt at analyzing the video game console industry by using social network approaches. We analyzed 14 game consoles in 5 generations and paid attention to the difference between the first and last. It has been shown that the industry is getting more crowded as each generation of console is introduced. The industry is settling on an average of 2 developers per publisher and 2 titles per developer. However, to be considered in the elite of any console, publishers must publish for 20 developers or developers must publish 20 titles.

As a future work, we may expand the analysis to include additional complex network metrics such as assortativity, betweenness and page rank centralities, and community detection including rich clubs. Moreover, the social network of the video game industry may be compared to other industries.

REFERENCES