Abstract

Massive Multiplayer Online Game (MMOG) has widely become a successful online business model. Current MMOG games lack the ability to stream dynamic user created audio signals, which makes an ultimate immersive experience beyond reach. In this paper, we propose a scalable immersive audio streaming platform which adds directional and spatial information to audio service of MMOG games. Our designs focus on providing a scalable platform that can stream dynamic audio content to a game involving millions of concurrent users.

1. Introduction

Massive Multiplayer Online Game (MMOG) has widely become a successful online business model. While having million of players online playing in the same shared world is every MMOG game developer’s dream, current game platform can only host concurrent users in the amount of thousands, not millions. While by using server clusters and dedicated server to divided game zones, current game can provide virtually a large shared world and many concurrent player, maintaining and configure these game clusters is so costly that prevent small and medium size games to deploy. Most MMOG developers opt to use server-client design to avoid game cheating and have better control over the ongoing game. Before a reliable distributed anti-fraud algorithm is viable, game developers will still have preference to centralized approach despite the benefits of distributed design, e.g., higher scalability and lower cost to deploy the system.

Sound, as an important component of entertainment business, is a key to immersive gaming experience. However, as current game becomes more interactive and dynamic, players are not satisfied with the pre-recorded sound effects stored locally. Players now want to talk to his/her teammates on a strategy shoot game or be able to play a sound track recorded by another player in a networked virtual environment. Currently this kind of audio streaming service, if offered in a game, is either done by forming a fixed audio group before the game starts, or by user initiated requests. The first approach works as an add-on feature to the game, can’t automatically provide game content related information. The second approach, although provide game related sound service to the game, usually doesn’t have a scalable design. None of these approaches is designed to provide an immersive audio experience to networked virtual environment.

In this paper, we propose an immersive audio streaming platform and designed to provide a scalable content-related sound solution to MMOG games. The concept can also applied to many other applications such as simulated training for soldiers and emergency workers.

We assume that the player/clients on the system can be trusted. The security issue, even we hesitate to put aside, is not discussed in this paper and is left to be an interest-
3.1. Virtual World Model

In this section, we will discuss the model of the virtual game world. Table 1 lists the definition of terms used in following discussion. In our design, the virtual world is a two-dimensional space with all objects (players and other game components) on the same plane. Every object has its own coordinate in the format of \((X, Y)\), with the most upper-left corner has \((0, 0)\) coordinate. By this definition, all objects on this world have positive coordinate readings.

Objects in the world can be either mobile or immobile. Mobile objects include players and all other objects that can be moved around during the game play. Immobile objects include object with fixed position, such as trees or houses. A Voronoi graph is calculated using a distributed sweepline algorithm [5]. Each node tracks its own coordinate during the game and update its neighbor nodes when necessary. To maintain a correct Voronoi graph, a peer-to-peer protocol similar to [6] is designed to cop with the dynamics of the voronoi diagram. Due to the space limit, we will focus our discussion to the audio deign.

3.2. Immersive Audio Design

In our design, the audio signals streamed from remote players (or objects in some cases) are rendered locally based from a local CD, to other players in an automatic and effective way. Scant attention has been paid to this streaming needs for massive online games and currently no major online game can provide integrated audio streaming service to its players except for simple voice chat which is usually initiated by users and is not an integrated experience of the game.

Table 1. List of terms used in this paper and their respective definitions.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>(S_i)</td>
<td>local audio signal captured at node (i)</td>
<td></td>
</tr>
<tr>
<td>(S_{i,j})</td>
<td>audio signal sent from node (i) and received at node (j)</td>
<td></td>
</tr>
<tr>
<td>(\theta_i)</td>
<td>direction angle of the audio signal at node (i)</td>
<td>(0 - 360^\circ)</td>
</tr>
<tr>
<td>(\theta_{i,j})</td>
<td>direction angle of the audio signal from node (i) to node (j)</td>
<td>(0 - 360^\circ)</td>
</tr>
<tr>
<td>(d_{i,j})</td>
<td>virtual distance from node (i) to node (j)</td>
<td></td>
</tr>
<tr>
<td>(f_{i,j})</td>
<td>sound fading parameter from node (i) to node (j)</td>
<td>(0 &lt; f_{i,j} \leq 1)</td>
</tr>
<tr>
<td>(I(X, Y))</td>
<td>coordinate of node (I) in the virtual world</td>
<td></td>
</tr>
<tr>
<td>(N(i))</td>
<td>set of all neighbor nodes immediately connected to node (i)</td>
<td></td>
</tr>
<tr>
<td>(Az_{i,80})</td>
<td>a virtual axis pointing to the east of the virtual world</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. List of terms used in this paper and their respective definitions.

ing future research topic. However, we claim that on most cases, e.g., in a simulated training environment or an online game system, sound is not an immediate target for fraudulence and a distributed solution is viable.

2. Related Work

Many designs have been proposed to use peer-to-peer topology for massive gaming environment. For example, in [7] authors proposed a application level protocol using Scribe [4] as the streaming platform. Also in [9] and [3] authors described a system design potentially can support up to millions of concurrent players.

Voronoi graph [8], which was well studied in geometry graphics before, recently has been implemented to construct peer-to-peer network. The latest research effort is presented in [6], which is using voronoi diagram to construct a peer-to-peer gaming platform.

However, to the best of our knowledge, there is no previous research that uses the voronoi diagram to create an distributed immersive audio platform for massive online games.

3. The Design

Sound, as an important component toward a immersive game experience, has been well studied and implemented in various types of games. Sound effect has been a area of intensive research since the beginning of computer gaming. Traditionally sound effects are stored locally and played from local depository during the game play. While this approach is good enough for many existing games, it is facing problems with the growing demand for interactive cooperation among multiple players. For example, in a massive online game, social activities are usually pervasive and audio exchanges among players, such as talks, have to be streamed from the remote players. This new requirement has created new challenges and opportunities for game designers: Design an immersive audio platform that can stream audio signals generated by players, e.g., voice chat or music played from a local CD, to other players in an automatic and effective way. Scant attention has been paid to this streaming needs for massive online games and currently no major online game can provide integrated audio streaming service to its players except for simple voice chat which is usually initiated by users and is not an integrated experience of the game.
on its directional and geographic characteristics. For example, a voice from a player nearby should sound louder than the voice from a player far away. Also the sound should be louder when two players are talking face to face. We call this audio experience, which renders the sound depends on physical features on the virtual worlds, immersive audio. The design presented in this paper, to the best of our knowledge, is the first peer-to-peer audio platform designed to provide immersive audio experience for massive online games.

Our algorithm used to calculate the Voronoi diagram and immersive sound effects is also very scalable. It runs in a local scale and thus will not become a performance bottleneck when the size of the players grows. The detailed math equations are discussed in Section 3.3.

At each node, mixing is performed to merge sound signals from neighbor nodes with the sound generated from local node. By mixing streams together, the bandwidth required for each audio stream is kept the same regardless of the total number of players in the system.

Our design also keep in mind the dynamics of the gaming environment and tries to cope with the fast changing game scenes with quick audio response. For example, in a first person view shooting game, when an enemy soldier approaches close enough, players will be able to hear the sound of foot steps even if the enemy is still out of the sight. These features will greatly increase the challenge and fun of the real time shooting games such as *Counter Strike* (CS).

### 3.3. Audio Mixing Model

In our voronoi diagram, a node $i$ is connected to its immediate neighbor node $j$ through a bi-directional link $L_{i,j}$. We denote all the neighbor nodes of node $i$ as $N(i)$. Each node $i$ can mix the audio streams received from its neighbor nodes, plus its local audio signal, if there is any, and
then forward the results to its neighbor nodes. This process is shown mathematically on Equation 1. We’d like to emphasize here that the result audio signals $S_{i,j}$ are usually different for each different destination node $j$.

$$S_{i,j} = [S_i \times \cos(\theta_{i,j}/2)] + \sum_{n=1}^{m} S_{m,i} \times \cos(\theta_{n,j}/2)] \times f_{i,j}^{m}$$

(1)

We denote a virtual axis that pointing to the east of the world as $Axix_0$. Then the direction of sound from a node $i$ can be represented by $\theta_i$, which denotes the angle between $Axix_0$ and the vertex indicating the peak of the sound magnitude (Fig 2.(a)). The magnitude of a sound signal generated at node $i$ is represented by $Si$. If a sound signal is multi-directional, e.g., the sound of an explosion, we can set the value of $\theta_{i,j}$ on Equation 1 to 0.

In order to add space characteristic to the audio service, we also calculate the distance between nodes. Each node in the virtual environment is given a x-y coordinate $(X,Y)$. As shown in Fig 2.(b), the distance $d_{i,j}$ between two neighbor nodes $i$ and $j$ is calculated as following equation:

$$d_{i,j} = \sqrt{(I_x - J_x)^2 + (I_y - J_y)^2}$$

(2)

We can calculate the magnitude of a sound on a certain direction by applying the original magnitude and the angles between them. Fig 2.(c) illustrated the angle between the direction of original sound generated locally at node $i$ and the destination node $j$. A similar sample using forwarding scenario is shown on Fig 2.(d).

### 3.4 Future Work

With the help of higher order voronoi graph, we can extend our design to support three dimensional game world in the future. Also, besides the directional and spacial parameters, additional features, such as echo, can also be added to the immersive audio design.

### 4 Performance Evaluation

In this section we will present the preliminary analysis of our design. As we mentioned before, the major contribution of our design is to provide a scalable immersive audio streaming platform for online games. We define immersive audio as directional and spacial sensitive audio experience. The directional information is preserved for each node $i$ as $\theta_i$ denotes the sound direction. And also the sound magnitude is adjusted based on the angle $\theta_{n,j}$ while forwarding audio stream from node $n$ to node $j$ through node $i$. By using Equation 1, a gradual audio forwarding chain is built. Nodes on the same direction of the sound will receive a strong audio signal while nodes on the opposite direction of the sound will receive very weak signal. This is not only to simulate what is happening in the real world, but also designed as a build-in echo canceller since the audio signal sent out from node $i$ can theoretically loop back to node $i$ again after as little as three hops. By applying the directional parameter, the original signal sent back to node $i$ will be minimized and ignorable.

The spacial information is used to simulate the fading of sound as it travels through the virtual space. The fading factor $f_{i,j}$ between node $i$ and $j$ ranges from 0 to 1. When it is set to 1, node $j$ will receive full strength of signal from node $i$ after the directional rules are applied. The smaller the fading factor is, the quicker the sound will fade while travels. If the fading factor is 0, e.g., sound proof door, no sound signal will be transferred directly from node $i$ to node $j$.

As a feature of Voronoi diagram, for a graph with $N$ nodes, the total amount of bits moved around in the network is bonded in $O(3N - 6)$. This shows that the network transportation resource is used in linear function with the size of the players.

Our design is also very scalable. All equations are calculated only using the information that can be retrieved from direct neighbor nodes. This guarantees that the execution of our system will not be slowed down even the number of participants is of thousands or even millions.

We do noticed that a voronoi graph can become very unbalanced in some cases such as shown in Fig. 3. However we advise that this situation can be avoid in a game world by applying the physical limitations to players, e.g., only certain number of players can surround another player due to the spaced taken by those players.

---

**Figure 3. Example of unbalanced Voronoi Graph**
5. PartyPeer: The Application

In order to show the viability of our design, we are implementing a RPG game called PartyPeer. This section will reveal some details of this work in progress.

PartyPeer is designed to provide a virtual party house for online social activities. Each player is represented as a virtual figure after login to the game. The game world is constructed by many virtual party houses that connected to each other. Each party house is a virtual space where all players inside form a voronoi graph and connected through the peer-to-peer topology. A player can approach other players in the same house and turn to the person who he want to talk to at the moment. The distance between nodes are used to calculated the sound fading effect and the directional angle is used to calculate the magnitude of sound in a certain direction. By applying both the directional and spacial information, an immersive audio experience is presented to players.

The PartyPeer is current designed to run on Windows operating systems. Torque game engine [2] is selected to implement the graphic front end, which looks similar to the graphic interface of a online social game called Habbo Hotel [1]. Fig. 4 shows the screen shot of a dance room in the game.

![Figure 4. Screen shot from Habbo Hotel](image)

6. Conclusions

In this paper, we presented a novel design that using peer-to-peer topology as streaming platform to provide immersive audio experiences to massive online games. We built our design on Voronoi graph and we added directional and spacial sound features as an integrated component of the game, which has not been done before. We briefly discussed the performance of our design and also presented the high level details of a massive online game called PartyPeer which runs on our proposed protocol. We conclude that our research is novel in the sense it is the first research approach on how to use peer-to-peer network to provide immersive audio environment to massive online games.

References