

Soluciones Visuales Interactivas aplicadas a Grandes Volúmenes de Datos de Entornos 3D de Aprendizaje y Prácticas

Juan Cruz Benito¹, Roberto Therón Sánchez¹ y Emilian Pizarro Lucas²

¹ Departamento de Informática y Automática, Universidad de Salamanca. Plaza de los Caídos s/n. 37008, Salamanca, España
{juancb,theron}@usal.es

² Servicio de Inserción Profesional, Prácticas y Empleo, Universidad de Salamanca. Patio de Escuelas 1. 37008, Salamanca, España
mili@usal.es

Resumen Este documento presenta el trabajo de investigación asociado al desarrollo de una propuesta de Visualización de Información apropiada para representar grandes cantidades de datos sobre movimientos de usuarios entre regiones de un Mundo Virtual de Aprendizaje y Prácticas y los patrones de exploración que en ellos se producen.

Este trabajo abarca todas las fases que incluyen en la propuesta de visualización: desde la identificación de las fuentes de datos, la extracción y procesado de la información, la aplicación de algoritmos de Minería de Datos para extraer conocimiento de la información, hasta la propuesta final de una solución visual que represente la información y el conocimiento que encierran los datos, y el conjunto de características y mecanismos que debe incorporar dicha visualización.

Keywords: Visualización de Información, Hierarchical Edge Bundles, HEBATT, Mundos Virtuales, Teleports, Minería de Datos

1. Introducción

Las tecnologías interactivas *online* o disponibles a través de Internet ofrecen la posibilidad de llegar a multitud de personas, sin importar la edad, cultura o nivel social. Una de esas tecnologías, es la conocida como Mundos Virtuales, o Entornos Virtuales Multi-Usuario (EVMU o MUVES). En un ellos, el usuario dispone de *alter ego* virtual o *avatar* que le representa en este entorno simulado. A través de estos *avatar*, al usuario se le proporcionan una serie de posibilidades de comunicación, colaboración, e interacción casi de cualquier tipo, tanto con el entorno virtual 3D como con el resto de usuarios que tienen presencia en dichas plataformas. Además, ofrecen un entorno *immersivo*, donde los usuarios experimentan una sensación de estar dentro de las situaciones y los eventos que ocurren, de un modo totalmente distinto a cómo se perciben en otras plataformas como la Web, u otros servicios informáticos en la Red.

El uso de este tipo de plataformas da lugar a la generación de grandes cantidades

de diversos tipos de datos, desde datos relacionados con la comunicación entre usuarios, a datos relacionados con los patrones de uso en objetos, o como es el caso del análisis, datos relacionados con el movimiento entre distintos territorios virtuales por parte de los usuarios. Debido al carácter inmersivo de la experiencia virtual de los usuarios, estos datos que se generan suelen ser indicadores fiables sobre distintos aspectos de la personalidad o comportamiento de los usuarios, ahí radica el interés del análisis de este tipo de entornos.

2. Objetivos

El objetivo principal de este trabajo de investigación es el estudio, comprensión y aprovechamiento de los datos sobre los movimientos de usuarios entre distintas regiones de un Mundo Virtual de Aprendizaje y Prácticas, para poder definir una visualización que permita a los distintos *stakeholders* implicados en el análisis de datos de un Mundo Virtual, descubrir conocimiento oculto en los grandes volúmenes de datos que se generan con información sobre los movimientos de usuarios entre distintos territorios virtuales, y que esta visualización pueda servir como Sistema de Soporte de Decisiones [1] sobre la actividad que realizan los usuarios dentro de un sistema 3D.

Además, se pretende dar un nuevo punto de vista sobre las posibilidades de visualización de datos extraídos de un Mundo Virtual, proponiendo un sistema bien construido, con un diseño atractivo, y que resulte útil y sencillo para el usuario, que abandone la línea típica de investigación de las visualizaciones de movimientos dentro de un Mundo Virtual [2] [3].

3. Fuentes y tratamiento de Datos

Para la obtención de datos de un Mundo Virtual de Aprendizaje y Prácticas, es necesario establecer conocer las fuentes de datos que se deben emplear en la extracción, y qué estrategia de extracción de información es conveniente aplicar [4] [5]. Una vez obtenidos los datos deseados del Mundo Virtual, es necesario procesarlos para extraer conocimiento de ellos. El procesamiento puede ser mediante técnicas clásicas, como puede ser la aplicación de procedimientos estadísticos simples, o mediante técnicas de minería de datos. En estas técnicas, destaca el algoritmo *Predictive Apriori*, el cual permite obtener patrones y reglas ocultas en los datos [6] [7], que más tarde se emplearán como herramienta de apoyo en las tareas de análisis, de modo que faciliten la comprensión del conjunto de datos y la obtención de conocimiento por parte del usuario.

El modelo de datos que se ha usado para el diseño de la visualización, ha sido extraído del Mundo Virtual de Aprendizaje y Prácticas USALSIM [8] [9].

4. *HEBATT: Hierarchical Edges Bundle Applied to Teleport Tracking*

Como se ha comentado previamente, el objetivo principal de este trabajo es proporcionar una visualización que permita analizar los movimientos entre territorios virtuales, y además aportando una solución novedosa y útil. Como resultado de este proceso de investigación, se ha llegado a la visualización denominada *HEBATT: Hierarchical Edge Bundles Applied to Teleport Tracking*. Como su propio nombre indica, se parte de la propuesta de Danny Holten sobre la visualización de datos jerárquicos *Hierarchical Edges Bundle* [10] [11]. En esta visualización de tipo radial, se codifica mediante variables visuales en forma de *arco* una jerarquía de datos, y se trazan líneas de tipo *spline* entre las distintas instancias de datos de dicha jerarquía que tengan relaciones entre sí [12].

Partiendo de esta idea resumida sobre el tipo de visualización *Hierarchical Edge Bundles*, se puede decir que en el conjunto de datos que se utiliza en el caso del Mundo Virtual, la jerarquía de datos se compone de una serie de islas o regiones, los grupos de usuarios que se integran en ellas, y los usuarios que pertenecen a esos grupos. Las líneas (curvas) que unen los distintos puntos de la estructura jerárquica representan los distintos movimientos que realizan los usuarios entre las distintas regiones.

La distribución de la jerarquía en el caso concreto de los datos extraídos del

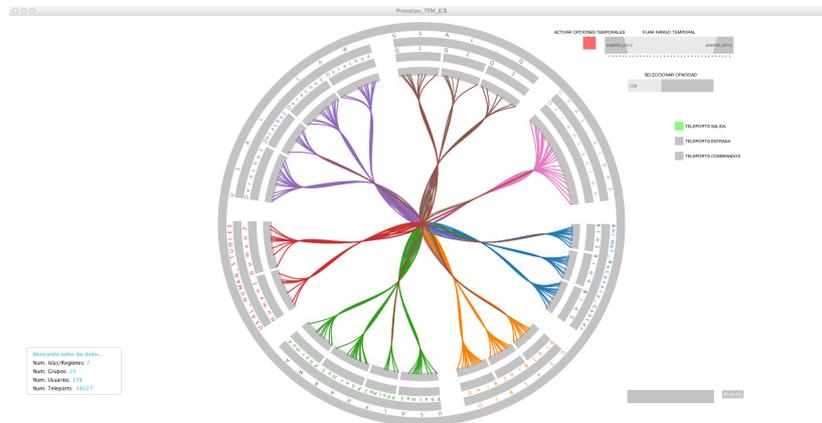


Figura 1. Vista general de la interfaz de la visualización HEBATT

Mundo Virtual USALSIM se realiza a través de cuatro niveles de arcos: un nivel superior que encierra toda la visualización y representa el conjunto completo, es decir, la vista general de todos los datos, de modo que si se selecciona en

cualquier momento ese arco exterior, se pasa a una vista general, como se ve en la figura 1. El nivel inmediatamente inferior a este que engloba todos los datos, es el que se corresponde con los distintos terrenos o islas que componen el Mundo Virtual; que en el caso del Mundo Virtual USALSIM, son siete territorios. Por debajo de ese nivel de la jerarquía y arcos de la representación visual, se presentan las distintas instancias de grupos de usuarios que existen en el Mundo Virtual. Cada una de las instancias de grupo se sitúa bajo el arco que representa la región o isla que sirve como base o localización de las actividades de aprendizaje y prácticas que realizan los usuarios de dicho grupo. Por último, el nivel más bajo de la jerarquía de datos, y nivel más bajo en la estructura radial, es para los arcos que representan a los distintos usuarios. En el caso de los usuarios hay un pequeño detalle a tener en cuenta: un usuario puede pertenecer a varios grupos, con lo cual, en una misma isla un usuario puede representarse varias veces, en función de la cantidad de grupos que pertenecen a esa región en los cuales se encuentra inscrito.

Como se observa en la figura 1, la interfaz de la aplicación cuenta con diversos elementos visuales que se organizan por toda la pantalla [13]. En la zona central se encuentra la visualización *HEBATT* propiamente dicha, la cual cambiará en función de la interacción que se de con el usuario. En la zona derecha de la interfaz se encuentran diversos elementos que permiten distintos tipos de interacción con la visualización, entre estos elementos se observan distintos grupos de botones, *sliders*, *sliders con funcionalidades de rango*, o un campo que permite introducir texto y realizar búsquedas en los datos. El uso de estos elementos se explicarán de forma más detallada en el apartado *Mecanismos de Interacción en la Visualización HEBATT* (4.9).

La zona izquierda de la visualización se ha reservado para la visualización de información textual relacionada con los datos que se estén mostrando dentro del *HEBATT*. Este texto se codificará mediante colores, que ayuden al usuario a la relación de los elementos visuales (mediante la característica de etiquetado, como se explicará en el punto 4.8).

En los siguientes sub-apartados se comentará qué tareas de análisis soporta la visualización y cómo es el comportamiento de la visualización en función de los datos que se quieran visualizar.

4.1. Tareas de Análisis que debe soportar la visualización *HEBATT*

Antes de mostrar el comportamiento de la visualización en cada caso concreto de uso, es preciso comentar qué tareas de análisis se permite a los usuarios:

- *Visualización general de los teleports realizados en cualquier instante de tiempo*
- *Visualización general de los teleports realizados en un instante de tiempo concreto*
- *Visualización de los teleports de salida de todas las regiones que componen el Mundo Virtual*
- *Visualización de los teleports de entrada a cualquiera de las regiones que componen el Mundo Virtual*

- *Visualización de los teleports de entrada y salida de cualquiera de las regiones*
- *Visualización de los teleports de salida de una región que realizan los usuarios inscritos en un grupo*
- *Visualización de los teleports de salida de una región que realiza un usuario concreto*
- *Visualización de los teleports de salida efectuados por los usuarios de una región concreta en un rango de tiempo seleccionado*
- *Visualización de los teleports de salida efectuados por los usuarios de un grupo concreto en un rango de tiempo seleccionado*
- *Visualización de los teleports que realiza un usuario en un rango de tiempo seleccionado*
- *Visualización de los patrones de teleports de salida de una región*
- *Visualización de los patrones de teleports de entrada de una región*
- *Visualización de los patrones de teleports de un usuario*
- *Visualización de los teleports que ha realizado un usuario durante su tiempo de uso del sistema, pudiendo recorrerlos a través de una escala temporal*

4.2. Tratamiento visual y lógico del volumen de datos a representar

En la propuesta de visualización *Hierarchical Edge Bundles (HEB)* original de Holten [10], ya se planteaban una serie de cuestiones que preparaban a la visualización para soportar grandes cantidades de datos, y se hacían diversas recomendaciones sobre cómo implementar la visualización. A continuación se comentan cuáles de las recomendaciones de Holten han sido utilizadas y qué otros mecanismos se han empleado en la visualización para gestionar grandes cantidades de datos:

1. **Uso de *splines*:** Holten, en sus publicaciones acerca de la visualización *HEB*, recomienda encarecidamente usar líneas que usen puntos de control y tensores para la formación de su curvatura, de esta forma es posible que se agrupen las líneas de forma efectiva y no saturen la visualización de líneas que no aporten información y oculten conocimiento. En el caso del prototipo software implementado, se ha hecho uso de líneas *splines* continuas [14], que usan tensores y puntos de control en su formación, aunque no permiten el control de estos tensores para la deformación de la curva en función de las necesidades que se tengan en la visualización.
2. **Uso de características de transparencia:** De la misma forma que en la característica anterior, Holten también recomienda el uso de características de transparencia, de forma que se pueda distinguir dentro de los conjuntos agrupados de líneas cuáles están compuestos por una cantidad mayor o menor de líneas (de relaciones en los datos), como se puede comprobar en la figura 2.
3. **Uso de líneas *resumen*:** Para la composición de la agrupación de los ejes hay dos opciones: una consiste en tener estructuras de datos que contengan cada teleport que se haya producido en el entorno virtual y toda su información asociada. La otra opción, es perder información como la referente a la

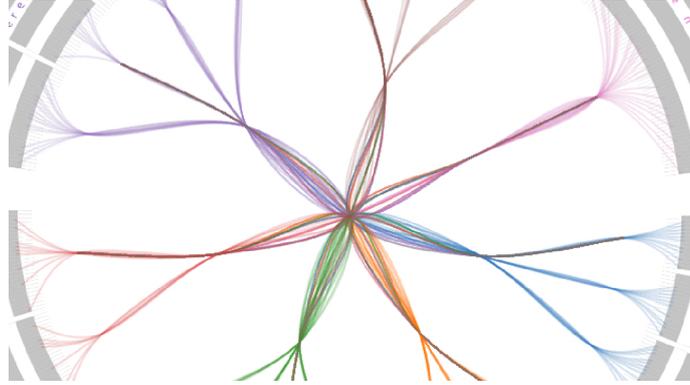


Figura 2. Transparencia aplicada en la agrupación de líneas de la visualización HE-BATT

fecha u hora de los teleports, y agrupar por cada tipo de teleport (definido por origen, destino y usuario), para así dibujar una sola línea *resumen*, que pueda expresar visualmente la cantidad de líneas que componen esta línea resumen. Esta representación visual se expresa también mediante la transparencia u opacidad de la línea, y esta opacidad se calcula de la siguiente forma:

$$\text{Opacidad Línea} = \text{factor Transparencia} * \text{num. de teleports del mismo tipo}$$

Por lo tanto, el resultado visual de estas líneas *resumen* es igual que en el caso de dibujar todas las líneas que componen cada agrupación, y la carga de trabajo que se realiza en cuanto al dibujado, es menor en una cantidad proporcional al número de líneas que no se dibujan.

4. **Uso de múltiples estructuras de datos:** Como se indica parcialmente en el punto anterior de dibujado de líneas *resumen*, para poder tener un rendimiento mejor de la visualización, es necesario implementar distintas estructuras de datos que ayuden a mejorar el rendimiento de la visualización, de modo que a través de la segmentación y agrupación conveniente de los datos se pueda abarcar el dibujado de grandes conjuntos de datos. En la implementación de este prototipo se han usado tanto estructuras *ArrayList* con conjuntos completos de teleports y teleports de tipo resumen, como estructuras *HashMap* para acceder a variables que se usan de forma recurrente de un modo más rápido que con los algoritmos clásicos de búsqueda.

4.3. La visualización *HEBATT* y la escala temporal

Como se ha comentado anteriormente, entre los datos que se obtienen en la extracción de teleports de los archivos log de OpenSim, se encuentra información relevante a fechas y horarios. Por lo tanto esta información puede (y *debe*) ser

utilizada para poder realizar visualizaciones de datos segmentados por características temporales. El autor MacEachren, en 1995 [15], planteaba una serie de preguntas que toda visualización que trabaje con datos debería poder responder. Estas preguntas básicamente ofrecen una guía para el análisis de datos con características temporales, y se encargaban, por ejemplo, de evaluar la existencia de un elemento en un momento dado, o cómo se distribuye la existencia de un objeto en el tiempo, qué tasas de cambios aparecían de los datos, o qué secuencias de cambios se daban, etc.

Tomando este conjunto de preguntas como guía de desarrollo, se han implementado varias características de manipulación de la visualización, de modo que se puedan segmentar los datos de la visualización en función de sus características temporales. Estas características son dos principales:

1. **Selector del rango temporal a mostrar:** A través de un selector de rangos (un elemento de interfaz similar a un *slider* pero con dos cabezas deslizantes), se ha implementado un sistema que permite visualizar los teleports de cualquier tipo en función del rango temporal seleccionado. Este selector se encuentra en la parte superior derecha de la interfaz. De esta forma se puede visualizar, en función del interés del analista, *qué conjuntos de teleports tienen lugar en el rango temporal seleccionado*, así como ir moviendo ese rango temporal en la escala del selector e ir viendo la *evolución temporal en los teleports* y extraer conclusiones sobre estas visualizaciones, como por ejemplo, *qué usuarios son los más activos, qué cantidad de teleports se asocian a un grupo en un tiempo concreto, en qué épocas hay una mayor carga movimientos o uso en el Mundo Virtual, si los teleports tienen características cíclicas en determinados usuarios*, etc. El rango temporal mínimo que se puede seleccionar es de un mes, el máximo dos años.
2. **Slider temporal de los teleports que realiza un usuario en el Mundo Virtual:** A través de la funcionalidad y el *slider* que se describe en la sección 4.7, es posible extraer información sobre todos los movimientos que ha realizado un usuario en el Mundo Virtual, estos movimientos se presentan ordenados temporalmente.

A través de este mecanismo se pueden realizar buena parte de las tareas de análisis que plantea MacEachren como, por ejemplo, *¿un usuario realiza movimientos cíclicos a lo largo del tiempo?, ¿Con qué frecuencia temporal realiza teleports el usuario?, ¿Existen cambios significativos en los teleports en función de la época del año?, ¿Qué cantidad de teleports realiza el usuario en cada época?*, etc.

4.4. Representación de teleports de salida

En este apartado se analizará el comportamiento de la visualización en el caso de la representación visual de los teleports de salida, sea cual sea el filtrado que se aplique sobre la visualización: *teleports de salida de una isla, teleports de salida relacionados con un grupo de usuarios, y teleports de salida relacionados con cada usuario concreto del Mundo Virtual*.



Figura 3. Visualización de los teleport de salida relacionados con una isla o región del Mundo Virtual

En la figura 3 se observa cómo se comporta la visualización en el caso de que se seleccione una región concreta y se quieran visualizar los teleports de salida. En la parte derecha de la interfaz, como se comentó anteriormente se encuentran una serie de elementos de control, y se observa que se ha seleccionado (recuadro en color verde) la visualización de los *Teleports Salida*.

En este caso (la visualización de los teleport de salida de una región) se observa cómo del conjunto de arcos de usuarios relacionados con los distintos grupos y una región concreta, surgen un conjunto de líneas curvas, codificadas con el color correspondiente a esa región (que se puede observar en el color de las etiquetas textuales de la región y los grupos que pertenezcan a ella). Estas líneas representan únicamente los teleport que han realizado los usuarios que pertenecen a grupos de aprendizaje y prácticas relacionados con esa región, no se observan los teleport de salida que han realizado otros usuarios que hayan pasado por la región, debido a que ese comportamiento no cumpliría los principios de la jerarquía de datos y la visualización jerárquica empleada.

En la parte derecha de la visualización, se observa una figura cuadrada roja acompañada de una inscripción que pone *Mostrar Patrones Teleport Región*. Este elemento, si se activa, muestra los patrones o reglas de movimiento que parten de esa región hacia otras, estos patrones de movimiento se han obtenido mediante los procesos de Minería de Datos (apartado 3), y se comentará su visualización en el apartado 4.6.

En el caso de los teleport de salida de un grupo, la visualización toma un aspecto parecido al caso de las regiones, aunque las líneas curvas que representan los distintos teleports surgen de los arcos que representan a los usuarios que pertenecen al grupo que se está analizando, estas líneas también estarán codificadas con el color correspondiente al grupo de usuarios, que coincide con el color asignado a

la región a la que pertenece el grupo. En el caso de este grupo de usuarios, se puede ver que se han movido a todas otras regiones del Mundo Virtual (incluso de su región a su región). Nótese que las líneas de llegada de un teleport a una región, se distinguen porque no llegan a tocar ningún arco perteneciente a una región, ya que el único que podrían tocar sería el de la región (un teleport no puede llegar a un grupo o a un usuario, sólo a un territorio), y si se estirase la línea hasta el arco de la región podría hacer perder parte de la simplicidad a la visualización, dificultando la comprensión de datos.

En el caso del análisis visual de los teleport de salida de un usuario (figura 4), el

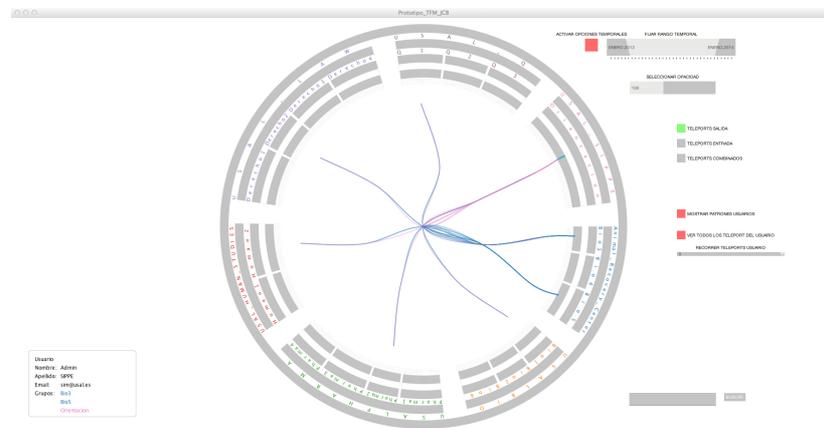


Figura 4. Visualización de los teleport de salida de las regiones a las que pertenece un usuario

comportamiento de la visualización es distinta, ya que muestra todos los teleport de salida de las distintas regiones a las que pertenece un usuario. Es decir, si un usuario pertenece a varias regiones, incluso a varios grupos por región, se mostrarán los teleport que salen de los arcos de ese usuario en cada región, viendo además a qué islas o regiones se dirige. A través de la codificación por colores en función de la región de donde salga el teleport, es posible distinguir a qué territorios se dirige el usuario en función de la isla en la que se encuentre. En este caso se observa que los elementos de control de la parte derecha de la interfaz cambian respecto a los otros casos. Cuando se analizan los teleport de un usuario se ofrecen dos opciones más, la visualización detallada de todos los teleport de un usuario, y la visualización de los patrones de teleport de dicho usuario; las visualizaciones que resultan de activar estas opciones se explicarán a continuación (apartados 4.6 y 4.7).

4.5. Representación de teleports de entrada

En el caso de que un usuario quiera analizar los teleports de entrada a una región, deberá seleccionar dicha opción en los elementos de control de la interfaz, que sólo funcionarán en el caso de que el usuario quiera conocer los teleports de entrada a cualquier región (teniendo activado el visionado de todos los datos), o en el caso de que esté evaluando una región concreta. En los casos del análisis de usuarios y grupos, la opción de visualizar los teleports de entrada no funcionará, debido a que, como se comentó anteriormente, no es posible llegar a un grupo o a un usuario, sólo se llega a un terreno.

Como se puede observar en el figura 5, cuando se selecciona una región y la

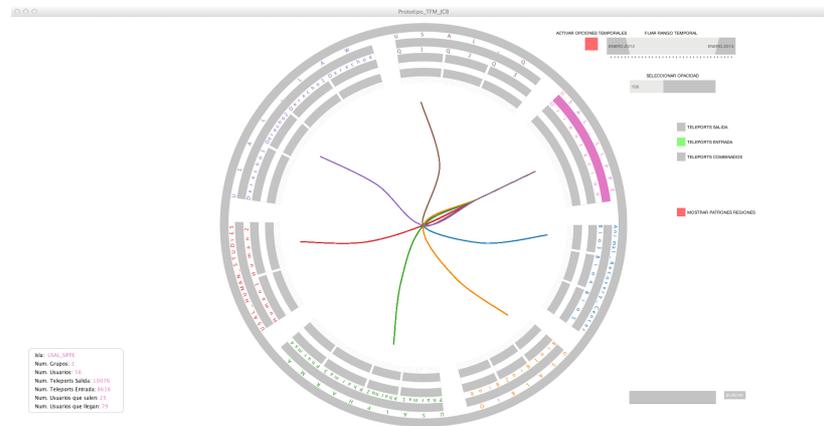


Figura 5. Visualización de los teleports de entrada a una región

opción de ver los teleports de entrada, se muestra una vista donde aparecen distintas líneas codificadas con los colores de las regiones de las que parten los teleports. Por lo tanto, a la región seleccionada llegarán una serie de líneas que parten de las distintas regiones del Mundo Virtual desde donde se haya llegado a la región seleccionada.

4.6. Visualización de los patrones de teleport

Como se apuntaba anteriormente, en las vistas de análisis de usuarios y regiones, se ofrece la opción de visualizar sus patrones de teleport, obtenidos mediante los algoritmos de Minería de Datos, por lo tanto ha sido necesario diseñar cómo un conocimiento extraído de esa forma podría ser representado [16].

En el caso de los patrones de teleport de usuarios, se ha optado por la representación a través de las líneas que representan el origen y destino de esos teleports. Estas líneas son de un color distinto al resto (color `#bcbd22`), y su ancho es

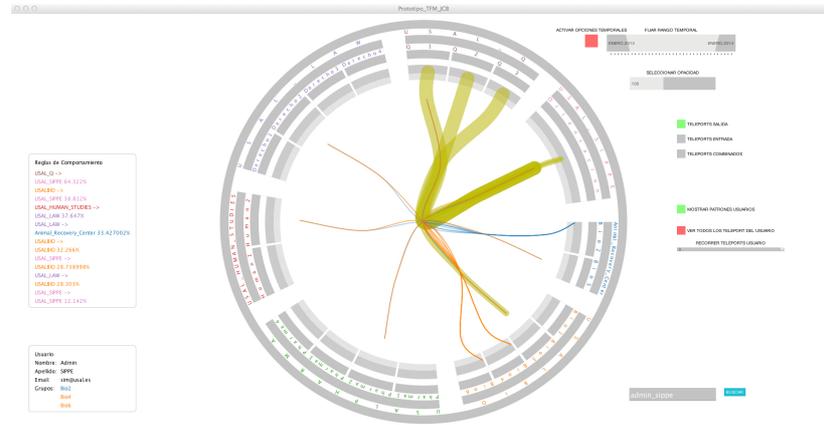


Figura 6. Visualización de los patrones de movimiento en el caso de los usuarios del Mundo Virtual

proporcional a la probabilidad de teleport entre las dos regiones que unen. De esta forma, cuando se representan visualmente los patrones de teleport, se muestran las líneas de teleport del usuario, de modo normal, y las que representan los patrones de movimiento, de modo que el contraste por el color y tamaño hacen que el usuario se fije rápidamente en ellas. Un ejemplo de esta representación visual se puede observar en la figura 6. Hay que señalar que en esta vista, sólo se representan, de igual modo que en los teleport de salida, los teleport de las islas a las que pertenece el usuario, no se pintan los teleport de las regiones en las que no está registrado el usuario. También es necesario apuntar que en la parte izquierda de la interfaz de usuario se mostrará la información textual acerca de las 10 reglas de comportamiento más relevantes para ese usuario, con el texto codificado como en otros casos con los colores de las regiones a las que hagan referencia.

En el caso de la visualización de los patrones de movimientos entre regiones, la representación visual se realiza de forma distinta. En este caso, cuando el analista elige ver los patrones de movimientos, desaparece de la visualización el conjunto de teleports que salen o entran a esa región seleccionada. En su lugar se dibuja un esquema de puntos y líneas curvas, de modo que cada punto, centrado en el arco de la región a la que se corresponde, se dibuja con una codificación en función del color de la región a la que representa. Una vez establecida la referencia entre los puntos correspondientes a las regiones del entorno virtual que estén presentes en las reglas o patrones de teleport, se dibujan las líneas que unen la región de origen con las de destino. En este caso las líneas se codificarán con el color de la región de origen, y representarán la probabilidad de ocurrencia de ese patrón de teleport a través de su ancho (figura 7). De igual forma que en los patrones de teleport referentes a los usuarios, en la parte izquierda de la pan-



Figura 7. Visualización de los patrones de movimiento entre regiones del Mundo Virtual

talla se mostrará la información textual sobre esas reglas, también codificando las regiones implicadas en cada regla con los colores de la región a la que hacen referencia.

4.7. Visualización detallada de los teleport de un usuario

En la vista de análisis sobre los teleport de los usuarios, se ofrece una opción de visualización detallada de todos los teleport que ha realizado un usuario a lo largo de la existencia de su *avatar* en el entorno 3D. Cuando se activa la opción *Ver todos los teleport del usuario*, disponible únicamente en la vista de teleports de un usuario concreto, se activa un *slider* en la interfaz, a través del cual, mediante el deslizamiento por la barra horizontal y su escala, es posible visualizar paso a paso y en orden cronológico todos los teleport que ha realizado un usuario. De este modo es posible ver la evolución *espacio-temporal* del usuario, pudiendo comprender en función de las fechas y el teleport que realiza, qué le mueve a realizar cada uno de ellos (para llegar a este nivel de descubrimiento de conocimiento, el usuario que analice estos conjuntos de teleport debe tener un amplio conocimiento previo de las actividades teórico-prácticas que existen y qué características tiene el entorno virtual). Como complemento a la visualización de líneas que van de una región a otra, se hace uso de información textual, aunque en este caso en la parte derecha de la interfaz, justo debajo del *slider* usado para hacer las transiciones sobre el conjunto de teleports. En esta información textual se incluye una descripción acerca del teleport que se está mostrando; con una descripción que incluye la región de origen, la región de destino, el usuario que la realiza, y la fecha y hora del teleport para que se comprendan perfectamente las características del teleport, y se pueda conocer si tiene relevancia o no en las tareas de análisis.

4.8. Codificación de colores

Como señalan múltiples autores, como Colin Ware [17] o Edwar Tufte [18], una correcta aplicación del color en la visualización puede hacer que una visualización de información sea más clara, más agradable para el usuario, y sobretodo más útil. Del mismo modo, Healey [19] sostenía que no se deben mostrar más de diez colores simultáneamente en la pantalla, ya que los usuarios dejan de diferenciarlos de forma clara y se pierde información. Por ello, se ha planteado una interfaz donde nunca se lleguen a visualizar más de diez colores, esto es posible mediante una cuidada organización en la distribución de elementos visuales y su asignación de colores.

En este trabajo de investigación, el color ha tenido una gran influencia, ya que



Figura 8. Vista detalle sobre la codificación del color en la información textual de cada tipo de elemento de la jerarquía de datos

se ha hecho uso de la capacidad de *Etiquetado* que posee (Ware [17]). Esta capacidad implica que un color es un buen recurso para asociar objetos visuales, y hace que el usuario pueda relacionar distintos atributos de objetos a simple vista y sin tener que realizar un complejo proceso mental de *emparejado* de entidades visuales. En cuanto a esta característica de etiquetado, cabe destacar que se ha aplicado de forma que un mismo color representa siempre la misma semántica en los datos, es decir, cada isla o región del entorno virtual, así como los grupos y usuarios que pertenecen a ella, tienen un único color de etiquetado, de modo que el usuario sea capaz de, sólo visualizando el color, relacionar el concepto con la isla, grupo o usuario del que se está mostrando la información (en la figura 8 se observa como incluso la información textual de cada tipo de datos se codifica por los colores correspondientes). Cada teleport que salga de la jerarquía de datos de una región, tendrá asociado un color, de modo que, por ejemplo, es posible distinguir la dirección de un teleport en función del color de la línea que representa el teleport.

En cuanto a la paleta o gama de colores disponibles en la visualización ([#1f77b4](#), [#ff7f0e](#), [#2ca02c](#), [#d62728](#), [#9467bd](#), [#8c564b](#), [#e377c2](#), [#c2c2c2](#), [#bcdd22](#),

#17becf), es necesario comentar que no es una escala determinada simplemente por características de diseño, sino que se trata de una escala desarrollada de tal forma que la diferencia entre las distintas tonalidades y colores que la conforman son lo suficientemente distintas como para que un usuario (sin problemas de visión del color) sea capaz de diferenciarlas sin problema. Esta escala ha sido propuesta por Mike Bostock y se conoce como *Categorical Colors* [20].

La aplicación de estos colores, a nivel de diseño concuerda con el objetivo de desarrollar una interfaz clara, sencilla y visualmente atractiva. Por ello se ha optado por el uso de colores claros para los datos que se muestren siempre, y colores vivos para centrar la atención sobre los datos que deben ser considerados de inmediato por el usuario o analista.

4.9. Mecanismos de Interacción en la Visualización *HEBATT*

A lo largo de este trabajo de investigación, se ha tenido especial cuidado con todos los aspectos que influyen en la calidad de la visualización, desde la codificación de variables visuales, la muestra de información [21], las herramientas de análisis que se permiten, etc. Dentro de este cuidado de los elementos que aseguran la calidad de la visualización, se encuentran unos muy especiales: *los Mecanismos de Interacción*. Estos mecanismos son métodos de interacción con la visualización, de modo que permiten realizar de una forma más sencilla las distintas tareas de análisis [22].

Entre los distintos Métodos de Interacción que se incluyen en la herramienta visual *HEBATT*, destacan:

- **Seleccionar:** Permitir seleccionar un elemento visual que se quiere analizar, hace que el usuario pueda generar visualizaciones de datos segmentadas que le facilitan la comprensión de la información que se le presenta. El método de interacción *seleccionar* es el que permite que cuando un usuario pinche en un arco de región, grupo o usuario, la visualización cambie para mostrar sólo los datos relacionado con el elemento sobre el que ha hecho clic.
- **Explorar:** El mecanismo de interacción *explorar*, permite al usuario modificar la visualización de modo que pueda cambiar los datos que desee sin mostrar el resto. Este mecanismo se implementa a través de los distintos elementos *deslizadores* o *sliders*, ya que en varios casos en los que se pueden aplicar (movimiento en la línea temporal de los teleports, y muestra sucesiva de los teleports de un usuario), mediante el movimiento del cabezal de este elemento se va modificando la vista, de modo que el usuario puede ir buscando la ocurrencia de datos que desee, sin que le molesten el resto de datos que no tengan relación con la selección que hace el usuario.
- **Filtrar:** Los filtros permiten al usuario seleccionar qué datos desea observar y cuáles no, ya que todos los datos que no cumplan el filtro seleccionado por el usuario no se mostrarán. En la visualización *HEBATT* hay distintos métodos de filtrado:
 - **Filtrados temporales:** Del mismo modo que se puede hacer un uso exploratorio del selector temporal, este sirve también para filtrar los datos de

movimientos en el espacio virtual en función de las fechas. Si se ajusta un rango de fechas, y se fija en algún punto de la línea temporal, se está realizando un filtrado por fecha, de modo que en la visualización sólo aparecerán los datos que se correspondan con ese rango de fechas definido.

– **Filtrados por tipo de teleport:** En el caso de la visualización de datos general y por regiones, es posible especificar qué tipo de teleport se quiere visualizar (entrada, salida o combinados). Esto se indica mediante un filtrado implementado a través de un elemento de tipo *radioButton*, que sólo permite que esté seleccionado un filtro por tipo de teleport en un instante concreto de tiempo.

– **Filtrados para patrones de comportamiento:** De forma similar al *filtrado por tipo de teleport*, en las vistas que en las que se analizan los datos de las regiones o lo usuarios, es posible seleccionar la opción *Mostrar Patrones (de usuario o región)*. En este caso este filtrado que permite mostrar o no los patrones de comportamiento es distinto al resto de filtrados, ya que en función del elemento donde se seleccione el filtrado, la vista de datos anterior desaparecerá o no. En el caso de los usuarios, los datos referentes al usuario seleccionado no desaparecen, sino que la visualización de los patrones de comportamiento *completa* la vista de los datos. Sin embargo, en el caso de la muestra de patrones de movimientos en las regiones, los teleport de la región se ocultan para mostrar únicamente la codificación visual de estos patrones de teleport entre regiones.

– **Filtrados mediante búsquedas:** El analista que utilice la herramienta de visualización, puede realizar filtrados mediante la introducción de texto en la caja de entrada textual habilitada para tal efecto. Mediante la correcta escritura del nombre de alguna región, grupo o usuario, el sistema filtrará la información que se está mostrando para pasar a representar visualmente sólo la información correspondiente al elemento que haya coincidido con la búsqueda.

- **Modificar la interfaz:** En este caso, al usuario se le permiten hacer modificaciones en los elementos visuales, cambiando así su aspecto para revelar si es posible datos que se encuentren ocultos o no sean claros. Esta modificación de la interfaz se realiza a través del *slider* “*Seleccionar Opacidad*” que permite al usuario seleccionar la opacidad o transparencia de las líneas que representan los conjuntos de teleports y así descubrir si hay relaciones en los datos que no se vean de un primer vistazo .

5. Conclusiones

En este trabajo de investigación se ha desarrollado un nuevo tipo de representación visual de los movimientos de los usuarios entre regiones de un Mundo Virtual de Aprendizaje y Prácticas, con lo cual se ha cumplido el objetivo que se planteaba inicialmente. En cualquier caso, no se puede asegurar que esta visualización sea totalmente válida, ya que no ha sido probada por usuarios reales, ni analistas de datos. Ese es un aspecto de los que se realizaría en futuras iteraciones en el proceso de investigación, ya que es fundamental validar la herramienta

mediante el uso, de modo que se pueda comprobar si los conceptos teóricos aplicados son lo suficientemente claros como para que un usuario del tipo que sea (experto, no iniciado, etc.) sea capaz de usar la visualización sin ayuda y con un tiempo de entrenamiento lo más corto posible.

Esta visualización se ha desarrollado para el uso concreto en el caso de los teleports en un Mundo Virtual, pero de igual forma se podría aplicar a cualquier tipo de movimiento migratorio, ya sea generado en un sistema digital o en la realidad. Por ejemplo, un esquema similar al que se presenta en este trabajo, se podría proponer para visualizar los flujos migratorios (de personas, de animales, etc.) entre regiones y países del mundo, ya que mientras se cumpla en parte la jerarquía de datos la visualización se puede adaptar a otros casos.

Bibliografía

- [1] J. P. Shim, M. Warkentin, J. F. Courtney, D. J. Power, R. Sharda, and C. Carlsson, "Past, present, and future of decision support technology," *Decision support systems*, vol. 33, no. 2, pp. 111–126, 2002.
- [2] K. Börner, W. R. Hazlewood, and S.-M. Lin, "Visualizing the spatial and temporal distribution of user interaction data collected in three-dimensional virtual worlds," in *Information Visualisation, 2002. Proceedings. Sixth International Conference on*, pp. 25–31, IEEE, 2002.
- [3] K. Börner and S. Penumarthy, "Social diffusion patterns in three-dimensional virtual worlds," *Information Visualization*, vol. 2, no. 3, pp. 182–198, 2003.
- [4] J. Cruz, R. Therón, E. Pizarro, and F. J. García-Peñalvo, "Análisis de datos en mundos virtuales educativos," in *Actas del XV Simposio Internacional de Tecnologías de la Información y las Comunicaciones en la Educación (SINTICE13)*, 2013. En Prensa.
- [5] J. Cruz, R. Therón, E. Pizarro, and F. J. García-Peñalvo, "Knowledge discovery in virtual worlds usage data: approaching web mining concepts to 3d virtual environments.," in *Proceedings Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support (Eureka-2013)*, 2013. En Prensa.
- [6] R. Agrawal, T. Imieliński, and A. Swami, "Mining association rules between sets of items in large databases," in *ACM SIGMOD Record*, vol. 22, pp. 207–216, ACM, 1993.
- [7] M. Kantardzic, *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons, 2011.
- [8] E. Pizarro Lucas and J. Cruz Benito, "Usalsim: Inserción profesional, prácticas y aprendizaje en un mundo virtual," in *Buenas Prácticas de Empleabilidad de Estudiantes Universitarios*, pp. 87–99, Universidad Miguel Hernández, Ministerio de Educación, Cultura y Deporte, 2012.
- [9] E. P. Lucas, J. C. Benito, and O. G. Gonzalo, "Usalsim: Learning and professional practicing in a 3d virtual world," in *2nd International Workshop on Evidence-based Technology Enhanced Learning*, pp. 75–82, Springer, 2013.
- [10] D. Holten, "Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 5, pp. 741–748, 2006.
- [11] D. H. Holten, *Visualization of graphs and trees for software analysis*. PhD thesis, Technische Universiteit Eindhoven, 2009.
- [12] B. Cornelissen, D. Holten, A. Zaidman, L. Moonen, J. J. Van Wijk, and A. Van Deursen, "Understanding execution traces using massive sequence and circular bundle views," in *Program Comprehension, 2007. ICPC'07. 15th IEEE International Conference on*, pp. 49–58, IEEE, 2007.

- [13] B. Shneiderman and C. Plaisant, *Designing The User Interface: Strategies for Effective Human-Computer Interaction*, 4/e (New Edition). Pearson Education India, 2003.
- [14] “Processing: Drawing curves.” <http://processing.org/tutorials/curves/>. R. E. (14). Julio, 2013.
- [15] A. MacEachren, *How Maps Work*. Guilford, New York, 1995.
- [16] D. A. Keim, “Information visualization and visual data mining,” *Visualization and Computer Graphics, IEEE Transactions on*, vol. 8, no. 1, pp. 1–8, 2002.
- [17] C. Ware, *Information Visualization: Perception for Design*. Elsevier Inc, 2013. ISBN: 978-0-12-381464-7.
- [18] E. R. Tufte, *Beautiful evidence*, vol. 23. Graphics Press Cheshire, CT, 2006.
- [19] C. G. Healey, “Choosing effective colours for data visualization,” in *Visualization’96. Proceedings.*, pp. 263–270, IEEE, 1996.
- [20] M. Bostock, “D3 categorical colors.” <https://github.com/mbostock/d3/wiki/Ordinal-Scales#categorical-colors>. R. E. (14). Julio, 2013.
- [21] B. Shneiderman, “The eyes have it: A task by data type taxonomy for information visualizations,” in *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pp. 336–343, IEEE, 1996.
- [22] R. Therón and L. Fontanillo, “Diachronic-information visualization in historical dictionaries,” *Information Visualization*, 2013. En Prensa.