**Instituto Tecnológico y de Estudios Superiores de Occidente**

Reconocimiento de validez oficial de estudios de nivel superior según acuerdo secretarial 15018, publicado en el Diario Oficial de la Federación el 29 de noviembre de 1976.

Departamento de Electrónica, Sistemas e Informática

Maestría en Sistemas Computacionales



**ANÁLISIS Y VISUALIZACIÓN DE GRAFOS USANDO REALIDAD VIRTUAL**

Tesis Que para obtener el grado de

Maestro en Sistemas Computacionales

Presenta: David Alfonso Velasco Sedano

Asesor: Dr. Luis Fernando Gutiérrez Preciado

Tlaquepaque, Jalisco. 1 de Noviembre de 2017.

RESUMEN

Los grafos son una representación visual de un conjunto de elementos dentro de un sistema. Se construye a partir de 2 elementos principales. Primero tenemos los llamados nodos, los cuales simbolizan miembros dentro del sistema (como un rúter dentro de un sistema de conexión de red tipo WAN). El segundo elemento se le denomina como arista, el cual representa una conexión directa entre 2 nodos (este puede ser a sí mismo en algunos casos).

Utilizando está representación podemos extraer información del sistema. Está siendo, y no limitada, a distancias y caminos entre diferentes elementos. Ver que tan robusto es el sistema y su conexión. Podemos entender el costo de atravesar ciertos caminos y cuáles son las rutas óptimas para alcanzar nuestro objetivo a partir de un punto inicial deseado.

Actualmente se encuentran herramientas (tal como Gephi) que permiten desplegar y hacer análisis de los grafos, aunque estos se encuentran representados en ambientes bidimensionales. Durante la investigación se va gestionar una solución para la visualización de un grafo en ambiente virtual tridimensional.

Al inicio el trabajo se estará enfocando a la solución únicamente para la red social llamada Twitter. Aunque se estará planeando tener un alcance también académico y para investigaciones. Generando rutas para poder alcanzar mejores recursos relacionados al tema de investigación, qué tan relevante es la investigación y colaboradores/autores trabajando en tópicos similares. Tomando como base de datos inicial el Repositorio Institucional del ITESO (REI).

TABLA DE CONTENIDO

[1. INTRODUCCIÓN 7](#_Toc497208574)

[1.1. Antecedentes 8](#_Toc497208575)

[1.2. Justificación 8](#_Toc497208576)

[1.3. Problema 8](#_Toc497208577)

[1.4. Objetivos 8](#_Toc497208578)

[1.4.1. Objetivo General: 8](#_Toc497208579)

[1.4.2. Objetivos Específicos: 9](#_Toc497208580)

[1.5. Mercado meta 9](#_Toc497208581)

[1.6. Hábitos de consumo actuales 9](#_Toc497208582)

[1.7. Características funcionales 10](#_Toc497208583)

[1.8. Novedad científica, tecnológica o aportación 10](#_Toc497208584)

[2. ESTADO DEL ARTE o de la TÉCNICA 11](#_Toc497208585)

[2.1. Visualización de redes en realidad virtual 12](#_Toc497208586)

[2.2. Un acercamiento inmersivo a la exploración visual geoespacial de base de datos de redes 12](#_Toc497208587)

[3. MARCO TEÓRICO/CONCEPTUAL 13](#_Toc497208588)

[3.1. Grafo 14](#_Toc497208589)

[3.2. Base de datos 14](#_Toc497208590)

[3.3. Framework 14](#_Toc497208591)

[4. Plan del proyecto TOG 15](#_Toc497208592)

[4.1. Alcance 16](#_Toc497208593)

[4.1.1. Estructura del desglose del trabajo (WBS) 16](#_Toc497208594)

[4.1.2. Requerimientos 16](#_Toc497208595)

[4.2. Cronograma 16](#_Toc497208596)

[4.3. Riesgos 16](#_Toc497208597)

[4.4. Secciones adicionales (Costos, calidad, recursos humanos, riesgos, comunicaciones). 16](#_Toc497208598)

[5. CONCLUSIONES 17](#_Toc497208599)

[5.1. Conclusiones 18](#_Toc497208600)

[5.2. Trabajo Futuro 18](#_Toc497208601)

LISTA DE FIGURAS

[Figura 4.1 Estructura del desglose del trabajo (WBS) 16](#_Toc497321448)

[Figura 4.2 Cronograma 18](#_Toc497321449)

LISTA DE TABLAS

[Tabla 4.1 18](#_Toc497321441)

LISTA DE ACRÓNIMOS Y ABREVIATURAS

|  |  |  |
| --- | --- | --- |
| REI |  | Repositorio Institucional del ITESO |
| ITESO |  | Instituto Tecnológico y de Estudios Superiores de Occidente |
|  |  |  |
|  |  |  |

INTRODUCCIÓN

En este capítulo se presenta brevemente los antecedentes del objeto de estudio, justificación del objeto de estudio, justificación y la definición del problema.

* 1. Antecedentes

El estado actual del arte con respecto al despliegue de cualquier grafo en un ambiente virtual tridimensional está limitado a proyectos estudiantiles. Mucho son prototipos que utilizan redes sociales como su fuente de información. Estos ligan al despliegue de los mensajes enviados por usuarios al igual que la creación de caminos para alcanzar a cualquier individuo en una red social.

También se encuentran herramientas de modelado de grafo y su interpretación. Aunque por el momento estas están limitadas a un ambiente bidimensional. Para mencionar algunas de estas, podemos encontrar a Grapheey y Gephi. Aunque estos carecen de un despliegue del grafo de forma tridimensional.

* 1. Justificación

Se debe hacer más eficiente la manera en que se consigue, analiza y se despliega información basado en grafos. Una representación de esto es en las redes sociales. Debe de generarse una forma de fácil acceso, digerir y entregar información relacionado a un usuario o un concepto.

* 1. Problema

Las redes sociales vistas como un grafo tienen consigo un gran desafío en su tiempo de ejecución y análisis. Esto a su vez conlleva al desafió de poder resolver la petición hecha por el usuario en un tiempo eficiente y entregar un resultado lo más cercano posible a lo deseado.

A la hora de hacer investigaciones (por ejemplo, aquellas relacionadas al área de mercadotecnia), se dispone de mucha información, aunque esta se entrega de una manera pobre y difícil de digerir. Lo cual conlleva a múltiples iteraciones de filtrado del contenido hasta obtener la información que se requería desde la primera consulta.

* 1. Objetivos
     1. Objetivo General:

Se estará desarrollando lo que se denomina un framework, el cual es una solución pre-hecha y lista para ser usada por otros desarrolladores. Esto funciona como una clase (al estilo de cualquier lenguaje orientado a objetos), donde simplemente tienes que crear un objeto de tipo de la solución y tendrás acceso a sus métodos públicos. Los métodos públicos estarán enfocados al filtrado y la generación de grafos tridimensionales. Aunque se puede extender a tener implementaciones abstractas las cuales el usuario final (como un investigador) puede definir para obtener una mejor respuesta basada en sus necesidades.

Dentro de nuestro escenario, se va construir el framework especializado en la red social Twitter. La cual debe tener la capacidad de generar los grafos tridimensionales en base a la meta-data de los comentarios, basado en las imágenes, entre otras opciones. Las opciones serán los requerimientos del usuario entregados a partir de un conjunto de filtros.

* + 1. Objetivos Específicos:

De forma colectiva se espera, y no está limitado, a los siguientes objetivos:

* Hacer una representación 3D del grafo usando datos obtenidos de la página Twitter. La complejidad se encuentra en la evaluación y lógica para distribuir en un espacio virtual los nodos de una manera óptima.
* Generar y gestionar filtros. Estos se deben representar como sub-grafos.
* Se debe disponer de una lógica de identificación de nodos principales. A su vez, estos se deben priorizar en su despliegue.
* Tener un sistema funcional de navegación del ambiente virtual donde el usuario pueda explorar.

De manera personal, se espera alcanzar por lo menos los siguientes objetivos:

* Expandir mi conocimiento de generación de ambientes tridimensionales usando Unity.
* Generar conocimiento de cómo proponer y crear algoritmos eficientes para solucionar problemas cuya respuesta tiene una representación visual.
* Ampliar mi conocimiento del análisis de grafos en el escenario de redes sociales, usando meta-datos como nodos y la probabilidad de éxito en tener un sub-siguiente meta-dato o imagen asociado a esté como arista.
  1. Mercado meta

El trabajo está dirigido a cualquier mercadólogo, sociólogo, psicólogo u otra persona que puede ser beneficiada con el análisis y compresión del comportamiento de redes sociales. Debe estar sujeto a que estará extrayendo información de manera masiva. Actualmente se cuenta con un conjunto de sociólogos y psicólogos fuertemente interesados en el tema.

En un futuro se planea que la aplicación tenga una especialización para el ambiente político. Siendo sus principales consumidores abogados, periodistas, investigadores, y cualquier persona interesada en este tema.

Por igual se desarrollará una extensión para personas dentro del gremio académico e investigadores. Donde el mayor enfoque sea la representación de trabajos relacionados a los temas que domina el usuario o que desea expandir.

* 1. Hábitos de consumo actuales

El estado actual del mercado en su mayoría está dominado por herramientas de visualización de grafos en dos dimensiones. Tal es el caso de las herramientas Centrifuge, Gephi, Graph-Tool y Graphviz. Muchas de estas herramientas piden que de manera manual que el usuario meta la información del cual el análisis se basará en.

Mientras tanto, podemos utilizar software similar para el análisis de una red social. Usando de ejemplo a GraphStream, Graph-tool, y AllegroGraph. Aunque siguen pidiendo que el usuario meta la información inicial para generar los grafos. Estas siguen estando limitadas en su mayoría en una visualización en dos dimensiones.

Finalmente, páginas de redes sociales tal como Facebook, Twitter y Youtube ya tienen soluciones que permiten el análisis de la red social. Su limitante es que no se despliegan como grafos al momento de escribir este documento. Sin mencionar que simplemente otorgar información de su propio sitio y no la relación entre plataformas.

* 1. Características funcionales

El sistema en su inicio va funcionar utilizando el dispositivo llamado HTC Vive. Este dispositivo permite interactuar y ver ambientes de realidad virtual. La solución se va dividir en 3 secciones principales.

Primero se encuentra la interfaz e interacción del usuario-máquina. En esta capa debemos ser capaces de interpretar las interacciones del usuario. Esto siendo cualquier entrada otorgada por él vía una interfaz gráfica (tal es el caso de los filtros deseados).

Después, en la segunda capa, se estará enfocando al despliegue del grafo mismo. Esto conlleva al cálculo de distancias “físicas” entre nodos. La forma en que vamos a colorear nodos. Al igual que aquí podemos crear las nubes de información relacionada.

Finalmente, nos vamos a encontrar con la capa del análisis y creación del grafo. Como el nombre indica, debe ser capaz de recibir los criterios obtenidos del usuario y encontrar la forma más eficiente de generar el grafo mismo. En ocasiones, deberá ser capaz de gestionar los sub-grafos y decir que nodos tienen mayor prioridad.

En esta investigación nos vamos a enfocar en la última sección. Se estará planteando distintos algoritmos y funcionalidades para la gestión de datos. Se busca hacer un comparativo entre diferentes soluciones propuestas con la nuestra. También haremos el esquemático de nuestro framework.

* 1. Novedad científica, tecnológica o aportación

Primero se estaría entregando una herramienta que su mayor enfoque es la visualización del grafo de manera virtual tridimensional. Donde el usuario pueda caminar entre y actualizar de acorde a sus necesidades.

Después tenemos la capacidad de generar el grafo en base a información dinámica que el usuario no mete de forma manual o que esté limitado a formatos. Esto ayuda a generar panoramas recientes y puntuales de cualquier tópico a analizar.

ESTADO DEL ARTE o de la TÉCNICA

En este capítulo se presenta un resumen de los trabajos relacionados con el análisis y visualización de grafos usando realidad virtual.

* 1. Visualización de redes en realidad virtual

En este documento se detalla la problemática de navegar en un ambiente virtual utilizando el teclado y ratón. Se propone como usando un equipo especializado para estos ambientes mitiga mucho el conflicto. Haciendo más intuitivo la navegación. [1]

* 1. Un acercamiento inmersivo a la exploración visual geoespacial de base de datos de redes

La investigación gira alrededor de las ventajas de ver la representación de un grafo vía un espacio virtual en vez del tradicional espacio bidimensional. Se menciona la curvatura y proyección de los elementos a desplegar. Haciendo referencia las distancias claves e ideales entre el usuario y el objeto a explorar. También se toca el tema de cómo el ojo va a ir observando y escaneando el área virtual. [2]

MARCO TEÓRICO/CONCEPTUAL

En este capítulo se presentan las bases teóricas y conceptuales sobre el análisis y visualización de grafos usando realidad virtual.

* 1. Grafo

Los grafos son una representación en forma de gráfica de problemas matemáticos. Estos tienden a representan objetos y lugares en la vida real, aunque también pueden representar conceptos abstractos. Se construye a partir de nodos y aristas.

Llamamos nodos a cualquier elemento dentro del sistema que estamos graficando. Pueden representar ciudades y personas hasta elementos como páginas web, rúters, entre otros. Los nodos se conectan a través de aristas. En algunos escenarios las conexiones pueden tener peso y/o dirección, de acuerdo sea el caso. [3]

* 1. Framework

Se considera un framework a cualquier solución de software que se puede re-usar y es considerado semi-completo. Se le denomina semi-completo a cualquier aplicación o implementación que todavía tiene métodos que se le necesita definir su funcionalidad.

Es esperado que un framework tenga una colección de métodos orientados a solucionar un solo problema. Aunque, por la definición de ser semi-completo, nos da un pequeño margen de cómo se debe solucionar el problema. [4]

Plan del proyecto TOG

## Alcance

### Estructura del desglose del trabajo (WBS)

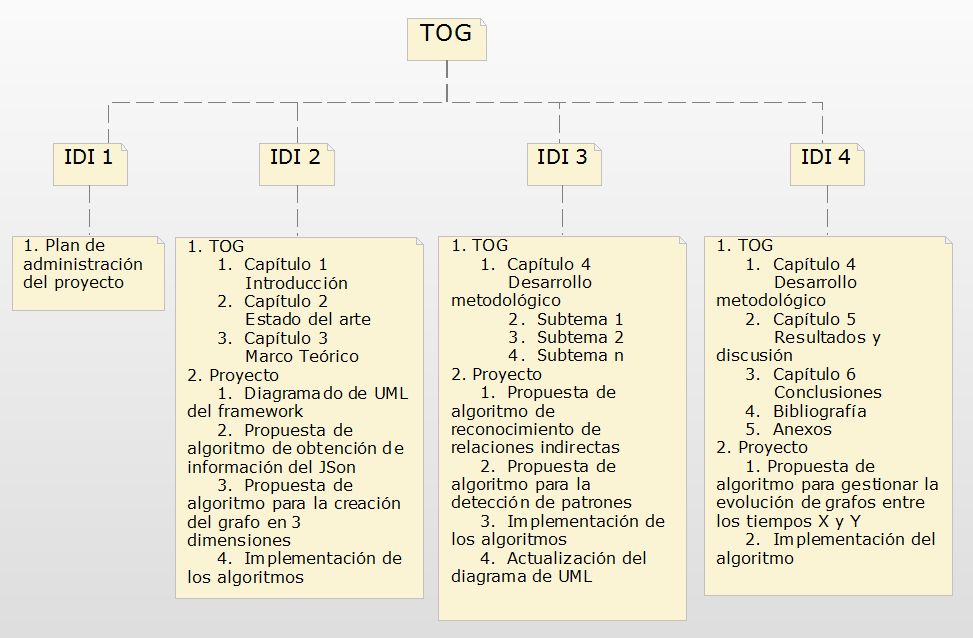


Figura 4. Estructura del desglose del trabajo (WBS)

### Requerimientos

|  |  |
| --- | --- |
| ID | Requerimientos |
| **IDI 1** | |
|
| REQ001 | Se debe considerar un año y medio de desarrollo |
| REQ002 | Debe incluir el WBS y Cronograma |
| **IDI 2** | |
|
| REQ003 | Las secciones provenientes de la hoja de concepto deben estar actualizadas y correctas en base a la retro-alimentación dada |
| REQ004 | Para la sección del estado del arte, se debe mencionar por lo menos tres distintas investigaciones y/o trabajos relacionados al análisis de grafos |
| REQ005 | En la sección de marco teórico, explicar por lo menos lo que es un grafo, lo que es un framework, cómo se hace una comparativa de algoritmos y qué es el patrón de diseño MVC |
| REQ006 | El diseño UML por lo menos debe incluir el diagrama de caso de usos. Diagrama de actividad para el proceso de convertir el archivo Json a un grafo y el proceso de filtrar la información. Debe haber el diagrama de clases de la estructura general |
| REQ007 | Cualquier propuesta hecha debe incluir el algoritmo o la solución que actualmente existe. Una comparativa entre los 2 y enseñar el porqué es más eficiente el algoritmo propuesto |
| REQ008 | Cualquier implementación de código debe seguir el patrón de diseño MVC. Toda función pública debe tener un valor de retorno, por default debe ser un booleano |
| REQ009 | Cualquier implementación de código debe ser revisado por el asesor por lo menos antes de ser incluido en el repositorio |
| **IDI 3** | |
|
| REQ010 | Todo tema expuesto en el desarrollo metodológico, en los casos de explicación de algoritmo, debe haber una explicación sobre sus duraciones y un análisis de su complejidad espacial y temporal |
| REQ011 | Los sub-temas puestos en el desarrollo metodológico deben reflejar las actividades principales del framework. Esto también incluye su evaluación de complejidad temporal y espacial |
| REQ012 | Para la propuesta de la relaciones indirectas, debe ser capaz de evaluar la relación de nodos de uno a uno y uno a múltiples |
| REQ013 | Para la evaluación de patrones, hacer un sistema que sea capaz de aprender y detectar los patrones de una manera automática |
| **IDI 4** | |
|
| REQ014 | Se espera que en resultados y discusión se pongan todas las comparativas entre el estado del arte y los algoritmos propuestos. Al final se tiene que poner los índices de mejora entre lo propuesto y lo actual |
| REQ015 | El algoritmo encargado de gestionar la evolución de grafos debe ser capaz de entregar múltiples estructuras (dependiendo de una frecuencia) que tengan funcionalidades que hagan la comparativa entre ellas sencilla |

Tabla 4.

## Cronograma

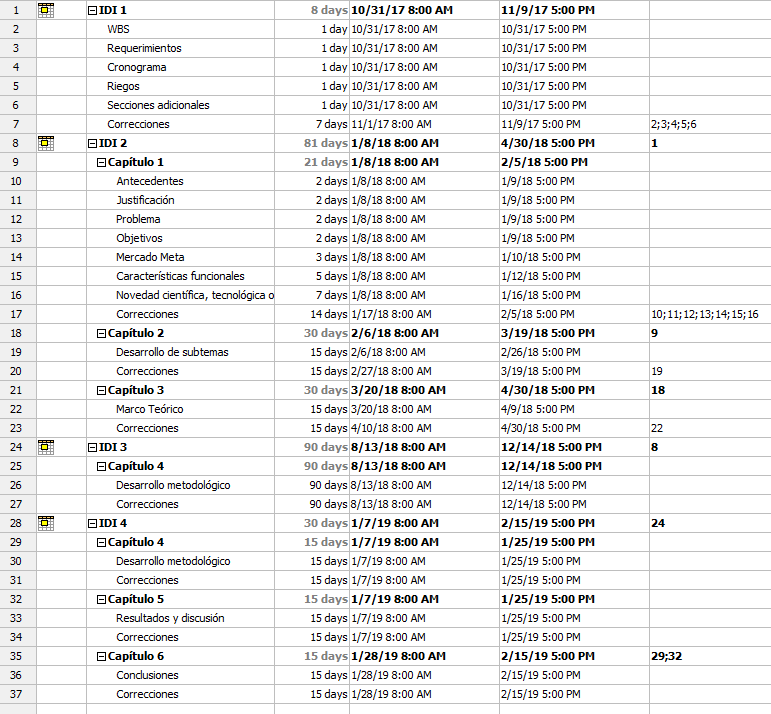


Figura 4. Cronograma

## Riesgos

## Secciones adicionales (Costos, calidad, recursos humanos, riesgos, comunicaciones).

# CONCLUSIONES

**Resumen:** [En este capítulo se presentan las conclusiones y trabajo futuro en relación a [el objeto de estudio]].

## Conclusiones

[Las conclusiones deben responde a los objetivos establecidos]

## Trabajo Futuro

[Se refiere a recomendaciones o descripciones sobre líneas de investigación que abre este trabajo, aplicaciones inmediatas que se derivan, o desarrollo de componentes o extensiones del desarrollo.]

BIBLIOGRAFÍA

|  |  |
| --- | --- |
| [1] | L. Beisel, P. Lamouric, J. C. Mayer y H. Dinh, «LUCA BEISEL,» Diciembre 2016. [En línea]. Available: http://lucabeisel.de/vr-network-visualization/. [Último acceso: 30 Octubre 2017]. |
| [2] | M.-J. Zhang y K. Zhang, «An Immersive Approach to the Visual Exploration of GeospatialNetwork Datasets,» de *CCS Concepts: Human-centered computing ~ Virtual reality;Human-centered computing ~ Geographic visualization*, Zhuhai, China, 2016. |
| [3] | D. P. Kumar, S. Archana y B. T. Kumar, «Graph Theory in an Object Oriented Approach,» *Journal of Computer Sciences and Applications,* vol. 3, nº 6, pp. 123-126, 2015. |
| [4] | D. C. Schmidt, «Applying Patterns and Frameworks to Develop,» de *The Handbook of Programming Languages*, St. Louis, Macmillan Technical Pub, 1997, p. 250. |

[Bibliografía Estilo IEEE: <http://www.ieee.org/documents/ieeecitationref.pdf>

El orden de la bibliografía debe ser el orden en el que aparece en el documento. Toda bibliografía puesta en esta sección deberá estar citada dentro del texto. La forma de hacer la cita cuando es una sola es [1], si son varias y salteadas es [3], [5], [8]. Si son varias consecutivas es [5]-[8]. Combinadas es [2]-[6], [9].

El formato de la bibliografía es como lista numerada simple, sin tabla.

El formato de cada una es como se describe a continuación y dependiendo del tipo:

*[Book Article*

* *For an article in an edited book, use practice similar to that for*[*author w/ editor or translator*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#auth-ed/trans)*above, inserting article title between author[s] names and book title.*
* *[Citation Number] Author name[s], "*[*article title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-art)*," in*[*book title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title)*,*[*editor names*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#auth-ed/trans)*, publication location: publisher, year,*[*pages.*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#pages)
* *Examples:]*

1. E.D. Lipson and B.D. Horowitz, "Photosensory reception and transduction," inSensory Receptors and Signal Transduction, J.L. Spudich and B.H. Satir, Eds. New York: Willey-Liss, 1991. pp. 1-64.
2. J. Lacan. "The insistence of the letter in the unconscious," in Psychoanalysis and Language, David Lodge, Ed., J. Rose, Trans., Ithaca, NY: Cornell University Press, 1992, pp. 123-34.

*[Journal*

* *[Citation Number] Author name[s], "*[*article title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-art)*,"*[*journal title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-per)*,*[*volume number, issue number, month (abbrv.)*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#volume)*,*[*pages*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#pages)*, publication year.*
* *Only include information which is pertinent to your source.  For example, many professional and academic journals do not have an issue month.  In that case, or when it seems unnecessary, do not include it in your citation.*
* *Examples:]*

1. K.A. Nelson, R.J. Dwayne Miller, D.R. Lutz, and M.D. Fayer, "Optical generation of turntable ultrasonic waves," Journal of Applied Physics, vol. 53, no. 2, Feb., pp. 1144-1149.
2. J. Allemang, "Social studies in gibberish," Quarterly Reviews of Doublespeak, vol. 20, no. 1, pp. 9-10.

*[Popular Periodical Article (monthly or bimonthly)*

* *In the case of popular monthly or bimonthly periodicals, omit volume number and issue, identifying instead by month and year of publication.*
* *[Citation Number] Author name[s], "*[*article title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-art)*,"*[*periodical  title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-per)*,*[*month (abbrv.)*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#volume)*,*[*pages*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#pages)*, publication year.*
* *Examples:]*

1. J. Fallows, "Network technology," Atlantic Monthly, Jul., pp. 34-36, 1994.

*[Popular Periodical Article (Biweekly, weekly, or daily)*

* *In the case of more frequently published periodicals, use day, month, and year to identify.*
* *[Citation Number] Author name[s], "*[*article title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-art)*,"*[*periodical  title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-per)*,*[*day number month (abbrv.)*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#volume)*,*[*pages*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#pages)*, publication year.*
* *Examples:]*

1. B. Metcalfe, "The numbers show how slowly the Internet runs today," Infoworld, 30 Sep., p. 34, 1996.
2. J. Turner, "Disorder 'kills without warning,'" The Toronto Star, 26 Jun., pp. F1-F2, 1998.

*[Paper Published in Conference Proceedings or Presented at Conference*

* *Treat a presentation in conference proceedings like an article in an edited book, including all available publication information.  Conference proceedings are often published by the organization holding the conference; in that case, do not cite the publisher.*
* *Example:]*

1. Paez-Borrallo, I.A. Perez-Alavarezz, and S.Z. Bello, "Adaptive foltering in data communications with self improved error reference," in Proc. IEEE ICASSP '94, 1994, pp. 65-68.

*[Treat an unpublished paper presented as a conference in the following manner: [Citation Number] Author name[s], "*[*article title*](http://www.edshare.soton.ac.uk/77/1/bbieee-help.html#title-art)*," presented at conference title. conference location, year.*

* *Example:]*

1. M. Lai, B. Chen, and S. Yuan, "Toward a new educational environment," presented at 4th Int. World Wide Web Conf. Boston, MA, 1995.

*[Web Page*

* *Give the author, title, type of medium (enclosed in brackets), volume and issue number (if on-line journal), page number (if relevant or given), and the year and the month of publication (in parentheses).  Then give the full internet address or the name of the online service provider prefaced by "Available at ".  If not an on-line journal, also put [cited year month day] before "Available at".*
* *Examples:]*

1. A. Harnack and G. Kleppinger, "Beyond the MLA Handbook: Documenting Electronic Sources on the Internet." Kairos, [Online serial] 1 (2), (1996 Sum), Available at HTTP: http://english.ttu.edu.kairos/1.2/
2. P. Curtis, "Mudding: Social Phenomena in text-based virtual realities," [Online document] Aug. 1992, [1996 Aug 30], Available at FTP: parcftp.xerox.com/pub/MOO/papers/DIAC921992.

**\*\*\*\*\*\*\*\*El estilo es Referencia**]

**APÉNDICE A. Título**