

Incidencia de la certificación, evaluación y temporalidad en la determinación de la confianza en la web semántica

Gustavo López¹ y Claudia León¹

¹ Universidad Central de Venezuela, Caracas Venezuela

RESUMEN

La confianza es un concepto complejo de representar matemáticamente, ya que está basada en factores sociales y psicológicos, siendo insuficiente evaluarla con sólo una variable de percepción, tal como se ha modelado hasta ahora.

La presente investigación plantea representar la confianza en la web semántica social estudiando la incidencia de variables hasta ahora no consideradas en conjunto, como lo son: la certificación, la evaluación y la temporalidad, con el fin de construir una red semántica social más confiable. Se usará y extenderá la ontología FOAF (friend of a friend) a objeto de proponer una nueva forma de representar la confianza.

Para probar la aplicación de esta representación se realizará un caso de estudio que consiste en crear una red de confianza de currículos haciendo uso de la web semántica social. Los individuos que formen parte de la red podrán evaluar y ser evaluados, así también podrán certificar información de terceros conocidos. Luego, se desarrollará un algoritmo de selección de personal que permita escoger al individuo adecuado para un cargo, basado en las competencias, el perfil y el grado de confianza que la red le otorgue.

Palabras Claves: Web Semántica, Web Semántica Social, Confianza.

ABSTRACT

Trust is a complex concept to represent mathematically, since it is based on social and psychological factors, making it hard to evaluate by only variables of perception.

In this paper we introduce a new approach to enable trust through certification, evaluation and time within the social semantic web realm. We will use and extend the FOAF ontology (friend of a friend) in order to propose a new way to represent trust.

To help prove this hypothesis, we introduce a set of tools that enable job matching using social semantic technology. The individuals that comprise the network will be able to evaluate and certify information over time. We extend current work in trust networks and create an RDF reasoner that, by combining a job-related trust network, will help provide us with the optimal person selection for a particular job.

Keywords: Semantic Web, Social Semantic Web, Trust.

1. Introducción

Internet se ha convertido en la principal herramienta de búsqueda e intercambio de información, sin embargo, existen inconvenientes a la hora de encontrar información precisa y confiable. Por ejemplo, si se quisiera buscar profesionales para emplear, lo común sería acudir a un buscador e introducir el perfil de los profesionales que se requieren, o en su defecto acudir a una página Web que ofrezca este servicio. Después de gastar mucho tiempo tratando de seleccionar entre toda la información arrojada con los candidatos más adecuados, se presentan las siguientes interrogantes ¿cómo confiar en que la información suministrada por el candidato es cierta? y más aún ¿Se puede confiar en que lo harán bien? Esta situación, es un problema de confianza.

La representación de la confianza no es un problema nuevo en el área de la web semántica, y en su solución se han desarrollado modelos y estrategias de confianza [KH04] [VS06] [DA07], funciones de creencia [LD04],

algoritmos de calidad de datos [DQ03], algoritmos de inferencia de confianza [GO07], entre otros. Algunas de estas propuestas utilizan teoría de grafos para asignar valores a la confianza a través de la apreciación [EW09]. Luego se aplican diferentes técnicas para inferir o propagar confianza en la red. El problema que se estudia en esta investigación es que la confianza es un modelo complejo de representar matemáticamente [EG98] ya que está basado en factores sociales y psicológicos, siendo insuficiente evaluarla con sólo una variable de percepción, tal como se ha modelado hasta ahora.

La presente investigación plantea evaluar la incidencia en la representación de la confianza, de factores hasta ahora no considerados en conjunto y valiéndose de la web semántica social como principal generador de información. Los factores claves a estudiar son:

- 1) Un proceso de certificación que permita confirmar por terceros la veracidad de hechos.
- 2) Una evaluación basada en apreciación ponderada, en la cual un conjunto de variables subjetivas arro-

jan un valor que mide el grado de confianza que tiene un individuo acerca de otro.

- 3) La temporalidad, ya que la percepción puede cambiar en el tiempo.

Como un ejemplo de la aplicación de esta nueva manera de representar la confianza se realizará un caso de estudio. El mismo consiste en crear una red de confianza de currículos haciendo uso de la web semántica social. Los profesionales que formen parte de la red podrán evaluar y ser evaluados, así también su información podrá ser certificada por terceros. Luego, se desarrollará un algoritmo de selección de personal que permita escoger el individuo para el cargo adecuado, basado en las competencias, el perfil y lo más relevante es que tomará en cuenta el grado de confianza que la red le otorgue.

Para hacer uso de la web semántica, es necesario que la información esté representada de forma que sea interpretada por los computadores, esté organizada y sea reutilizable. Para que esto ocurra, es necesario que la información se codifique mediante ontologías. El presente trabajo usa y extiende la ontología FOAF (friend of a friend) [FOAF], acrónimo de amigo de un amigo, que describe personas, sus actividades y las relaciones con otras personas u objetos.

La extensión que se plantea realizar a la ontología, se basa en las nuevas variables de información mencionadas anteriormente: certificación, evaluación y temporalidad.

Esta ontología no solo permitirá mejorar la búsqueda de información en el proceso de selección de personal, ya que la información seguirá un esquema común, sino además permitirá realizar inferencias de los mejores candidatos basado en la red social de confianza del usuario que consulta. Por su lado la ventaja de utilizar la red social será la de permitir obtener información útil basada en contribuciones e interacciones colectivas.

En la sección dos, se especifica el objetivo principal que se persigue en la investigación así como los objetivos secundarios que se plantean para satisfacer de manera metódica el cumplimiento del objetivo principal. En el entendido de que la representación de la confianza es un tema amplio de investigación, también, en esta sección, se establece hasta dónde se llegará y las limitaciones del estudio, asimismo se describe la importancia de la investigación, el aporte a la disciplina y el impacto actual de la confianza en las redes sociales.

Por otra parte, se ha mencionado que la representación de la confianza en la web semántica ha tenido gran atención en años anteriores, es por ello, que en la sección tres se referencian los trabajos relacionados en el área.

En la sección cuatro se explica la investigación a realizar, igualmente se informa el estado actual de la investigación así como los resultados esperados.

Y por último, la sección cinco presenta las conclusiones de este estudio.

2. Planteamiento del problema

2.1 Objetivo Principal

Representar la confianza de la web semántica social considerando las variables certificación, evaluación y tiempo para la construcción de una red social semántica social confiable.

2.2 Objetivos Secundarios

- 1) Estudiar la confianza en la web semántica social en un contexto de selección de personal para comprobar la incidencia de la introducción de las variables certificación, evaluación y tiempo en la confianza.

- 2) Extender la ontología FOAF (friend of a friend), acrónimo de amigo de un amigo, con la inclusión de las variables tiempo, factores de certificación y evaluación para crear modelo de representación de la confianza en la web semántica social.

- 3) Desarrollo de una red social que permita evaluar y certificar en el tiempo una red de currículos semánticos.

- 4) Crear algoritmo de búsqueda (Top K) haciendo uso de la ontología extendida para inferir en el campo de selección de personal.

- 5) Con el fin de demostrar la incidencia de la certificación, evaluación y temporalidad en la determinación de la confianza en la web semántica se realizaran pruebas.

2.3 Alcance

El alcance de esta investigación se circunscribe en la evaluación de la incidencia de las variables certificación, evaluación y temporalidad en la representación de la confianza, valiéndose de la web semántica social como principal generador de información.

Para demostrar la incidencia de estas variables se creará un sistema de red social que permita evaluar y certificar en el tiempo una red de currículos semánticos con el fin de realizar selección de personal, haciendo uso de la web semántica social en donde dado un cargo necesario se pueda inferir quien es la persona más adecuada para desempeñarlo.

Asimismo se desarrollará un algoritmo de selección Top K, focalizado en la estructuración de los datos semánticamente, en el uso de los mismos y en la creación de una fórmula que evalúe el candidato, según la confianza de su red social, basada en la certificación, evaluación y temporalidad.

Las optimizaciones a nivel de rendimiento, paralelismo y seguridad del algoritmo Top K no serán contempladas en esta investigación, ya que no es el objeto de estudio.

Las pruebas del algoritmo y la red social serán realizadas con datos ficticios de currículos, relaciones y evaluaciones. Se simularán servicios de certificación de entidades certificadoras como universidades y empresas. Dado un cargo, el

algoritmo de confianza arrojará un resultado con el candidato recomendado, este resultado se comparará con un proceso manual y controlado de selección de talentos realizado por personas, lo cual permitirá concluir los resultados del estudio.

Esta investigación no pretende demostrar hipótesis de comportamiento social.

2.4 Justificación

¿Cómo confiar en que la información encontrada en Internet es certera? Es aquí donde entra el tema de la confianza y la importancia que tiene en el ámbito de la Web. La representación de la confianza es un modelo matemático para ser aplicado en computación, que ha tenido atención en los últimos años y puede ser aplicado a muchos ámbitos, sin embargo existen factores hasta ahora no considerados para su modelado, los cuales se plantean introducir en este trabajo, valiéndose de la Web semántica social como principal generador de información.

Además de aportar en el desarrollo de la representación de la confianza, esta investigación también plantea contribuir en el área de la web semántica y en la formalización de la capa de confianza. Actualmente, los buscadores Web permiten realizar búsquedas de información mediante palabras claves que aparecen en el código HTML de las páginas Web. Los resultados que se arrojan en muchos casos no son útiles. Si por ejemplo, en un proceso de selección de personal se quiere buscar en la Web a un “ingeniero”, que se encuentre en una ciudad “Maracaibo”, y que tenga “experiencia laboral en petróleo”, la forma actual de buscar es introducir en el buscador las palabras claves: ingeniero, Maracaibo, experiencia laboral petróleo. La información que seguramente encontrará el buscador puede estar relacionada con demandas laborales, cursos, buscadores de empleo, entre otros y en el mejor de los casos candidatos a empleo. Normalmente se gasta mucho tiempo en seleccionar la información que puede ser útil. A través del caso de estudio se avanza en el uso y extensión de la web semántica en aras de facilitar la búsqueda de información en Internet.

Otras opciones de búsqueda utilizadas actualmente para reclutamiento y búsqueda de empleo son las redes sociales de profesionales. LinkedIn es una de las más populares. Una de las características de LinkedIn es que permite publicar datos como experiencia, educación, páginas web y recomendaciones, además de que permite establecer contacto con otros miembros enfocados a un ámbito profesional específico. Aunque LinkedIn permite obtener recomendaciones tipo texto que escriben contactos de la red, no contempla procesos de certificación de información, medición por evaluación de desempeño y temporalidad, variables que permitan confiar en que la información que yace en el currículo es certera o que permita escoger un candidato de acuerdo a la evaluación de su red social. El tipo de recomendaciones que dispone LinkedIn son escritas, lo que supone que si se requieren evaluar 10.000 CV también se tienen que evaluar las recomendaciones de cada uno. ¿Cómo medir cual candidato tiene mejores recomendaciones? En este sentido, esta investigación aportará la inclu-

sión de un algoritmo Top K de selección basado en variables de confianza.

Así como el área de selección de personal, existen otras áreas de aplicación, donde una representación de la confianza en redes sociales aplicaría para apoyo en la toma de decisiones y mejorar la búsqueda y confiabilidad de la información que yace en Internet, por ejemplo selección de compañías, productos, marcas, médicos, entre muchos otros objetos.

3. Antecedentes / Estado del Arte

En la web cualquiera puede decir cualquier cosa sobre cualquier cosa. Es por ello que cuando un usuario busca cierta información en la web aplica su juicio para descartar y llegar hasta lo que quiere. Pero este juicio, hecho por las personas es complejo ya que pondera inconscientemente gran cantidad de conocimiento acerca del mundo en que vivimos, experiencias pasadas personales que al final se resumen en una reputación percibida acerca de la fuente y la información que está suministrando. Esto no es posible en la web semántica donde el contenido debe ser reflejado usando ontologías y axiomas, dado que no sólo son las personas quienes consumen la información. Los agentes necesitan una manera de calcular automáticamente estos juicios de confianza para seleccionar un servicio o fuente de información.

Modelos y estrategias de confianza [KH04] [VS06] [DA07], funciones de creencia [LD04], algoritmos de calidad de datos [DQ03], algoritmos de inferencia de confianza [GO07], entre otros, han sido desarrollados. Muchas de estas propuestas utilizan teoría de grafos para asignar valores a la confianza a través de la apreciación. Luego se aplican diferentes técnicas para inferir o propagar confianza en la red.

A continuación se enumeran diferentes investigaciones realizadas sobre la confianza en la web semántica agrupada en categorías que responden a la manera como establecen en cada caso la confianza entre pares:

3.1 Confianza basada en política

La utilización de políticas para establecer la confianza se basa en manejar e intercambiar credenciales. Trabajos en confianza a través de políticas asumen que la confianza es establecida obteniendo suficiente nivel de credenciales de un ente específico y luego se aplica una política para permitir el acceso a esta información. Generalmente se utiliza un tercero con credibilidad para verificar y otorgar confianza, así como el uso de firmas digitales.

Entre los trabajos realizados se pueden citar PolicyMaker, Keynote y REFEREE que utilizan este concepto de verificar credenciales y establecer relaciones de confianza entre cada ente con la principal meta de establecer un método de control de acceso [VL05]. Por ende, su concepto de confianza es limitado a verificar las credenciales y restringir acceso a recursos dependiendo de políticas definidas por su aplicación. Por ejemplo, el dueño de una información provee acceso a otro si éste puede verificar sus credenciales

directa o indirectamente. Esto resuelve el problema de autorización y control del acceso.

3.2 Confianza basada en reputación

En este método se utiliza la reputación de un ente para establecer la confianza basándose en interacciones pasadas, las cuales son combinadas en fórmulas para predecir comportamiento futuro. Investigaciones en este método utilizan la historia de una entidad (actividad/comportamiento) para calcular confianza y pueden utilizar la confianza basada en referencias (información de otras personas), cuando no exista información de primera mano. Un ejemplo de esto son trabajos que utilizan las recomendaciones para generar confianza. eBay, WebOfTrust, IMDB, SPORAS, HISTOS, XREP, NICE, DCRC/CORC, Beta, EigenTrust establecen enlaces de confianza entre sus pares (entes) y asignan un valor de confianza a las relaciones [IS06]. El valor de esta confianza es una función donde se combina la percepción global y la evaluación basada en la percepción individual. En EigenTrust [SK03] (algoritmo utilizado en sistemas P2P para bajar música) consideran confianza como una función de archivos corruptos vs buenos. Un peer (en español: par) se califica como más confiable (por sus otros peers) si éste proporciona archivos buenos, por ende asignándole un peso de mayor confianza cada vez que sigue interactuando de buena manera. El cálculo de la confianza en EigenTrust es una variante del algoritmo de PageRank [PS98]. Un peer realiza un cálculo de la confianza basado en su rendimiento histórico.

3.3 Confianza basada en contexto

La confianza basada en el contexto se basa en utilizar metadatos en circunstancias donde la información acerca de algo ha sido afirmada. Por lo general, se analiza quién lo dijo, por qué lo dijo, dónde y cuándo.

Trabajos como el de TRELIS [RY02] abordan la confianza en dos términos, contenido y quién lo dice (la fuente). Describen un método para calcular confianza acerca de la fuente dependiendo de lo que dice la fuente. A medida que los usuarios dicen algo, ellos pueden incluir estadísticas de credibilidad y responsabilidad por los otros. Estos datos después son ponderados.

3.4 Confianza basada en mecanismo de contenido

Este método no utiliza metadatos acerca de la información sino reglas y axiomas en conjunto con el contenido de la información. Ejemplo de políticas siguiendo este método son: Confía en información que ha sido citada por lo menos en cinco fuentes independientes. Por ejemplo, si una tesis ha sido citada en más de cinco trabajos, se puede confiar en ese contenido. Los mecanismos de confianza basados en contenido y contexto no requieren evaluaciones de terceros, sino hacen uso de información existente en la web.

3.5 Confianza basada en redes sociales

Existe una nueva clasificación para asignar confianza, se basa en analizar las relaciones que juegan los actores dentro de una red social y sacar conclusiones acerca de la reputación utilizando análisis de redes (un campo relacionado a teoría de redes). Por ejemplo Regret [JS01] identifica a grupos en una red social, y NodeRanking [JP02] identifica a expertos dentro de una red social.

Algunas de estas metodologías terminan representando la confianza a través de grafos donde los nodos son personas y los arcos son valores para representar la confianza, sin embargo no toman en cuenta el tiempo. Tomando como referencia el algoritmo TidalTust [JG06] utilizado en la aplicación Web FilmTrust [JF06] pregunta la percepción de una película a una persona, sin considerar el factor tiempo, ya que se toma como una opinión estática. Por ejemplo, si María evalúa en el 2009 la película E.T. que vio en el año 1982, quizás no sea la misma respuesta que se obtendría si se le hubiese preguntado inmediatamente después de haberla visto.

En el caso de LinkedIn, una de las redes sociales de profesionales más populares actualmente, la cual permite publicar datos como experiencia, educación, páginas web y recomendaciones, además de poder establecer contacto con otros miembros enfocados a un ámbito profesional específico. Cuando se crea un vínculo en LinkedIn, un profesional X invita a un profesional Y a estar conectados.

En términos de grafos lo que se está haciendo es creando un arco punteado dirigido de Y a X (ver figura 1) que indica que está a la espera de que se certifique que Y conoce a X, que tienen un tipo de relación (colega, compañero, negocios, amigo, familiar, otro) y de dónde se conocen (compañía, institución educativa, otro).

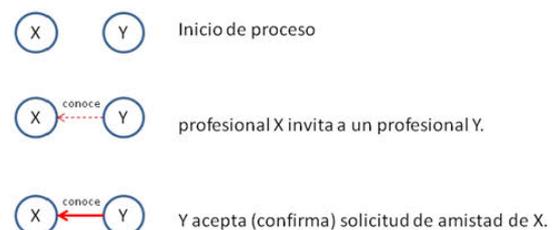


Figura 1. Confirmación de solicitud en LinkedIn

Cuando el profesional Y acepta la conexión se crea un arco sólido, que certifica que se conocen. Ver figura 1. Se observa en primer lugar que no existe un arco de X a Y, en segundo lugar que el arco certifica que el profesional Y conoce a X, pero no certifica que la información del currículo de X es cierta, tampoco permite que Y evalúe a X como profesional y viceversa.

4. Estado de la investigación

De acuerdo al alcance y los objetivos planteados, esta investigación ha avanzado con interesantes aportes en la

representación de la confianza y la evolución de la web semántica social, que a continuación se describen:

4.1 Arquitectura de Certificación

Las distintas categorías de investigaciones de confianza en la web semántica, definidas en la sección 3, tales como: política, reputación, contexto, y contenido permitieron diseñar una arquitectura de certificación.

Fundamentándose en que la confianza se basa en reputación (Ver antecedentes), la certificación de un tercero con credibilidad es una variable imprescindible para representar la confianza. Para este caso de estudio las universidades y empresas serán los entes certificadores de hechos que comprueben que un individuo estudió o trabajó en un periodo específico. Un colega o amigo también es un ente certificador pero que tiene menos peso o reputación, pero que vale si no se tiene certificación de la fuente principal. Así entonces para iniciar la representación de una red de confianza se creó una arquitectura que permite certificar hechos estáticos en el tiempo, tales como: lugar de trabajo, lugar de estudio, área de trabajo, área de estudio, tiempo de trabajo y estudio. La arquitectura de certificación en la red de confianza para el caso de estudio se muestra en la figura 2:

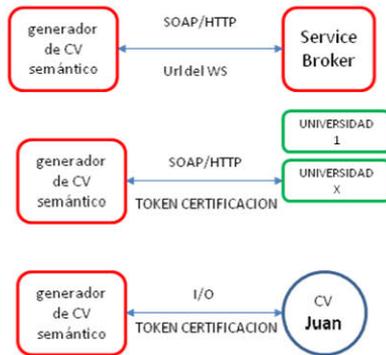


Figura 2. Arquitectura de certificación para el caso de estudio

Así pues, una arquitectura general de certificación, se ilustra a continuación:

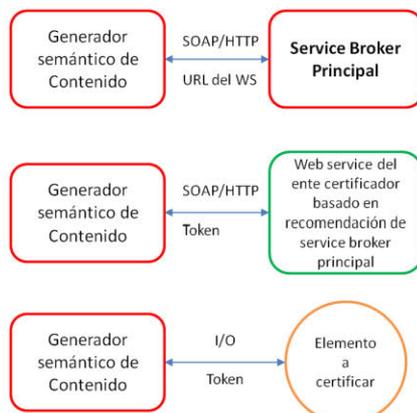


Figura 3. Arquitectura generalizada de certificación

Para darle aplicabilidad a esta arquitectura se realizó un “generador de CV semántico” que maneja un conjunto de servicios web. El primer servicio actúa como un “service broker” o un servicio que tiene los registros de ubicación de los demás servicios web que son los que certifican si la información de trabajo o estudio es verdadera.

Una vez ubicado el servicio web certificador, éste “certifica” la información dentro de un currículo. Una vez certificado la información se tiene un token en el propio CV para verificar su validez en cualquier momento.

Por ejemplo, supongamos que Juan en su currículo muestra que estudió su licenciatura en computación en la Universidad X. Lo primero que hace el “generador de CV semántico” es preguntarle al “service broker” donde puede validar “estudios en la Universidad X”. Este le da una dirección del servicio web de la Universidad X. Luego el “generador de CV semántico” solicita la certificación de esta información. En caso de que el servicio web de la Universidad X valide exitosamente la información de que Juan estudió en la Universidad X su licenciatura en computación, entrega al “generador de CV semántico” una token etiquetado para certificar validez del mismo. Este token permite invocar en cualquier momento el servicio web de la Universidad X y entrega el mismo (así como los otros datos suministrados) re-certifica que esta información es verdadera.

En caso de que no exista ente certificador principal, por ejemplo si se trata de una compañía pequeña que no tiene servicio de certificación, también se tomará en cuenta la confianza basada en mecanismo de reputación, muchas personas diciendo un verdad se acerca a la verdad. Siguiendo el mismo ejemplo de la compañía pequeña, los colegas pueden certificar que el profesional trabajó en esa empresa. De esta manera se tiene una certificación, aunque de menor peso.

4.2 Grafo de Evaluación

Fundamentándose en que la confianza se basa en contexto (Ver antecedentes) ¿qué dijo y quién lo dijo? fueron variables tomadas en cuenta para la representación de la confianza en este estudio. Por ejemplo, los candidatos para ser supervisores estarán evaluados por sus redes sociales de supervisados positiva o negativamente, esta evaluación tendrá un peso más alto que aquellas evaluaciones realizadas por sus amigos, familiares o relacionados que si bien pueden evaluarlos no tienen suficiente credibilidad porque no han tenido experiencia directa de trabajo.

Sea $G=(V,E)$ un grafo completo dirigido denominado “grafo de evaluación” con un conjunto de vértices V que representan a un individuo realizando trabajo u educación en un tiempo X , y un conjunto E de aristas, que son pares ordenados de elementos de V que representan la evaluación que tienen un individuo sobre otro por el trabajo realizado en este tiempo X .

Visto en forma de grafo, un ejemplo de evaluación se representa en la figura 4:

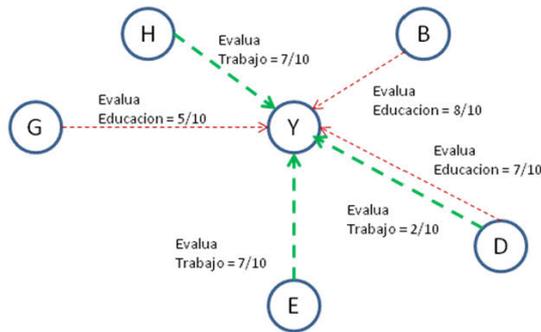


Figura 4. Ejemplo de Grafo de Evaluación (G) para caso de estudio

Para demostrar la aplicabilidad de este grafo se desarrolló un sistema de red social que permite a profesionales publicar los datos de sus currículos, con la facilidad de evaluar y ser evaluados por su red social, así como la certificación de la información por fuentes certificadoras (simuladas). La evaluación consiste en una encuesta 360 de valoración, compuesta por preguntas aleatorias, asignadas dependiendo el tipo de relación, en donde se puede evaluar a los contactos según una escala de valores dados. También se desarrolló un algoritmo Top K basado en el mecanismo de contenido (Ver antecedentes), el cual es un conjunto de reglas y axiomas con la información de los currículos. Este algoritmo permite realizar selección de personal, basado en la certificación, evaluación y temporalidad.

4.3 Grafo de Temporalidad en la certificación y evaluación

Fundamentándose en la categoría de investigación de confianza en la web semántica basada en reputación, se tomaron los conceptos de temporalidad en la certificación y evaluación. Dado que la variable tiempo juega un papel importante en la definición de confianza se decidió crear un grafo que la tome en consideración.

Un ejemplo que puede ilustrar el concepto de temporalidad es el siguiente: Supóngase un trabajo A donde solo dos actores María (M) y Juan (J) son protagonistas. Ver figura 5.

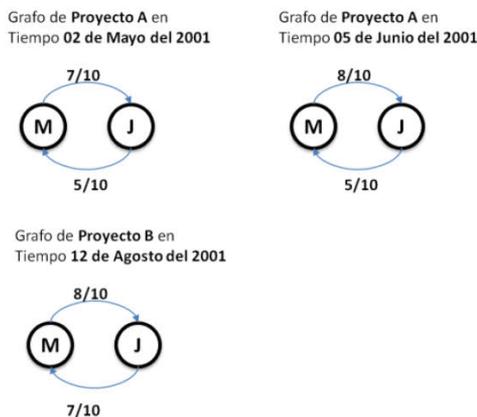


Figura 5. Grafo de evaluación en el tiempo para caso de estudio

Se puede apreciar que en el grafo 02 de Mayo de 2001 la apreciación de M hacia J es de 7/10 y de J a M es de 5/10. Estos resultados provienen de la ponderación de la evaluación entre M y J. El mismo caso aplica para la certificación.

4.4 Extensión de la Ontología FOAF

La utilización de ontologías permitió representar conceptos del dominio laboral y las relaciones entre estos conceptos. Esto a su vez ayudó a razonar y deducir las propiedades del dominio. Existen muchas ontologías para anotar a individuos en la Web semántica social tales como FOAF [FOAF], microformats [MICR], HR-XML [HRXM] y RDF Resume schema [RO07].

Se extendió la ontología FOAF. FOAF (Friend Of A Friend) es una ontología dentro de la web semántica para describir relaciones y actividades entre personas y objetos utilizando RDF (Resource Description Framework).

Para la extensión de la ontología FOAF se creó una clase llamada "Certificación". Esta clase permite incluir información adicional, que no se encuentra en FOAF, complementando la subclase Persona, con propiedades tales como: fecha de la certificación, ente que realiza la certificación, token que aprueba esta certificación, y el web service utilizado. Esta clase permite certificar el historial de trabajo, la educación y los cursos de una persona. A continuación se muestra la extensión realizada:

```

Class: foafExt:Certificacion
Certificacion - Una certificación.
Status: stable
Properties include: fecha, enteCertificador, token, webservice-
Certificador
Used with: hasWorkHistory hasEducacion hasSkill
Subclass Of: Persona
La clase certificación representa un token que un hecho fue certifi-
cado. Para recertificar el hecho se debe ir directamente al
webservice.
    
```

De la misma manera se creó una clase llamada "Evaluación". Esta clase permite incluir información adicional, que no se encuentra en FOAF, y también complementando la subclase Persona, con propiedades tales como: fecha que se realiza la evaluación y token de la evaluación. Dado que la evaluación es confidencial, éstas estarán almacenadas en un repositorio, el cual será utilizado únicamente por el buscador Top K. Aquí pueden existir cuantas evaluaciones se necesitan o quieran, dado que las evaluaciones no pueden ser vistas ni cambiadas.

Esta clase permitirá evaluar historial de trabajo, educación y habilidades.

```

Class: foafExt:Evaluacion
Certificacion - Una evaluacion.
Status: stable
Properties include: fecha, token
Used with: hasWorkHistory hasEducacion hasSkill
Subclass Of: Persona
La clase evaluación representa un token de una opinión acerca del
estudio, trabajo o habilidad. Las evaluaciones no pueden ser vistas
ni cambiadas. Estas están almacenadas en un webservice utilizado
únicamente por el buscador Top K, motivado a la confidenciali-
dad.
    
```

4.5 Herramientas de pruebas

Para probar la incidencia de la evaluación, certificación y temporalidad en la representación de la confianza, se desarrollaron dos herramientas y se realizaron diversas pruebas, las cuales están explicadas al final de esta sección.

Red social semántica de talentos:

La primera herramienta es una red social que permite certificar y evaluar en el tiempo una red de currículos semánticos. Ver figura 6:



Figura 6. Red Social de Talentos

Buscador social semántico:

La segunda herramienta es un “buscador social semántico”. Una vez que la “red de confianza” está realizada, utilizando un algoritmo de selección, el buscador tiene como objetivo obtener los mejores candidatos al cargo adecuado. Este algoritmo de “selección” tiene elementos basados en la clasificación de la confianza según contexto, contenido, política, reputación y redes sociales, tomando en cuenta la certificación, evaluación y temporalidad para buscar los mejores k candidatos a un puesto.

A continuación se muestra la fórmula para la selección de personal:

$$R(p_i) = \sum_{p_j \in M(p_i)} PR(p_j) * E(p_j) * RE(p_j) * T(p_j) * L(p_j) + \sum_{p_i \in N(p_i)} I(p_i) + \sum_{p_b \in O(p_i)} PR(p_b) * RE(p_b) * T(p_b) * L(p_b) * P(p_b)$$

Donde p_1, p_2, \dots, p_N son las personas bajo consideración, $M(p_i)$ son las personas que tiene evaluaciones sobre p_i , $PR(p_j)$ es el page rank de esta persona, $E(p_j)$ es la evaluación ponderada de la encuesta realizada por esta persona, $RE(p_j)$ es el tipo de relación con la persona, $T(p_j)$ es la ponderación del tiempo cuando se hizo la evaluación, y $L(p_j)$ es la distancia ponderada de lugares físicos mientras se hacía la evaluación. Esta primera sumatoria corresponde a la evaluación en el tiempo.

También $N(p_i)$ son las certificaciones que se tienen sobre p_i . Esta es la sumatoria que corresponde a la certificación realizadas por instituciones.

También, $O(p_b)$ son las personas que tiene certificación sobre p_b , $PR(p_b)$ es el page rank de esta persona, $RE(p_b)$ es el tipo de relación con la persona, $T(p_b)$ es la ponderación del tiempo cuando se hizo la certificación, $L(p_b)$ es la distancia ponderada de lugares físicos mientras se hacía la certificación y $P(p_b)$ es la ponderación de certificaciones

realizadas por institución vs las realizadas por personas. Esta es la sumatoria que corresponde a la certificación realizadas por personas en el tiempo.

Casos de pruebas:

Se realizarán 4 conjuntos de pruebas, en las cuales dado un cargo X (Ejemplo: Líder de proyectos de electrónica), se debe seleccionar el candidato mas adecuado. Los curriculums se tendran tanto en papel como en el sistema de la red social. Primero se realizará un proceso manual a través de la evaluación de curriculos en papel realizada por una persona. Por otra parte se correrá el algoritmo Top K . Se compararan ambos resultados para su posterior análisis y conclusión. Para ello se han creado 4 escenarios:

- 1) Escenario 1: Currículo con información falsa. Se tienen 3 currículos de los cuales uno es el mejor pero éste tiene información falsa.
- 2) Escenario 2: Habilidad para escribir currículos. Se tienen tres currículos de los cuales uno está incompleto pero éste es el mejor evaluado en la red social.
- 3) Escenario 3: Profesional de confianza. Se tienen tres currículos iguales. Sin embargo uno es un profesional conocido directamente por la persona que evalúa el currículo, otro tiene una relación indirecta y otro es desconocido.
- 4) Escenario 4: Múltiples variables: Se tienen tres currículos iguales. Uno es evaluado excelente por su red social, otro es bueno y otro es malo. Uno es desconocido, otro guarda una relación indirecta y otro es un conocido respectivamente.

Resultados esperados:

Con este tipo de pruebas se espera los siguientes resultados:

- 1) En cada prueba se espera que el buscador semántico seleccione de forma instantánea al candidato más adecuado basado en análisis de información del currículo persé y de información proveniente de la red social.
- 2) Comprobar que la certificación, la evaluación y la temporalidad son variables que inciden en tener la confianza de que una información es certera.
- 3) Comprobar que las redes sociales tienen gran potencial en la búsqueda de la confianza de la información.

5. Conclusiones

Durante este trabajo de investigación se evaluó la incidencia en la representación de la confianza de variables hasta ahora no estudiadas en conjunto, valiéndose de la web semántica social como principal generador de información. En este sentido, se llegó a las siguientes conclusiones:

- 1) La certificación, evaluación y la temporalidad son variables importantes construir una red social semántica más confiable.
- 2) Las redes sociales permiten generar la información necesaria para probar los modelos de confianza.
- 3) Extendiendo la ontología FOAF es posible que los datos sean almacenados en un formato distribuido, no controlados por un ente central, para que éstos puedan ser analizados por cualquier buscador.
- 4) Las redes sociales de profesionales actuales no cuentan con facilidades de certificación y evaluación en el tiempo. Este estudio aporta en primer lugar un nuevo modelo que puede ser utilizado para la selección de personal basado en la medición de la confianza de que la información suministrada es certera, en segundo lugar este mismo modelo puede ser utilizado para otro tipo de redes sociales.

References

- [KH04] Kieron O'Hara, Harith Alani, Yannis Kalfoglou, and Nigel Shadbolt. Trust Strategies for the Semantic Web. Workshop on Trust, Security, and Reputation on the Semantic Web. Hiroshima, Japan : ISWC2004, November 2004. <http://eprints.ecs.soton.ac.uk/10028/01/ISWC04-OHara-final.pdf>.
- [VS06] Trust Policies for Semantic Web Repositories. Vinicius da Silva Almendra, Daniel Schwabe. s.l. : 2nd Semantic Web Policy Workshop (SWPW'06), 2006. http://www.l3s.de/%7Eolmedilla/events/2006/SWPW06/programme/paper_14.pdf.
- [DA07] A survey of trust in computer science and the Semantic Web. Donovan Artz, Yolanda Gil. June 2007. Web Semantics: Science, Services and Agents on the World Wide Web. Vol. 5 Issue 2, págs. Pages 58-71. 1570-8268.
- [LD04] Weaving the Web of Belief into the Semantic Web. Li Ding, Timothy Finin. 2004. WWW2004. <http://ebiquity.umbc.edu/v2.1/get/a/publication/74.pdf>.
- [DQ03] Data Quality on the Web. .4 Report about a Dagstuhl seminar trying to relate information quality and the Web. <http://sirius.cs.ucdavis.edu/Dagstuhl03/dq.pdf>.
- [GO07] SUNNY: A New Algorithm for Trust Inference in Social Networks, using Probabilistic Confidence Models. Golbeck, Ugur Kuter and Jennifer. Vancouver, British Columbia : s.n., July 2007. Proceedings of the Twenty-Second National Conference on Artificial Intelligence (AAAI-07). www.cs.umd.edu/~golbeck/downloads/Sunny.pdf.
- [EW09] Wang, Eric. A Survey of Web-based Social Network Trust. Sydney, Australia : Department of Computing Macquarie University, 2009. http://web.science.mq.edu.au/~rdale/teaching/itec810/2009H1/WorkshopPapers/Wang_Eric_FinalWorkshopPaper20090523.pdf.
- [EG98] Gerck, E. Toward Real-World Models of Trust: Reliance on Received Information. The Meta-Certificate Group (MCG). [En línea] 1998. <http://mcwg.org/mcg-mirror/trustdef.htm>.
- [FOAF] Introducing FOAF. The Friend of a Friend (FOAF) project. [En línea] <http://www.foaf-project.org/original-intro>.
- [BL99] Berners-Lee, Tim y Fischetti, Mark. Weaving the Web. Harper San Francisco : s.n., 1999. Chapter 12. 9780062515872.
- [SEMw] Semantic Web. Wikipedia. [En línea] http://en.wikipedia.org/wiki/Semantic_Web.
- [XML] Extensible Markup Language (XML). W3C. [En línea] <http://www.w3.org/XML/>.
- [XMLS] XML Schema. W3C. [En línea] <http://www.w3.org/XML/Schema>.
- [RDF] RDF. W3C. [En línea] <http://www.w3.org/RDF/>.
- [RDFS] RDF Vocabulary Description Language 1.0: RDF Schema. W3C. [En línea] <http://www.w3.org/TR/rdf-schema/>.
- [OWL] OWL Web Ontology Language . W3C. [En línea] 2004. <http://www.w3.org/TR/owl-semantics/>.
- [TG93] T., Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Knowledge Systems Laboratory, Stanford University. CA : s.n., 1993. págs. KSL- 93-04.
- [TA93] Tello, Adolfo Lozano. Ontologías en la Web Semántica. Área de Lenguajes y Sistemas Informáticos. Departamento de Informática. Escuela Politécnica. , Universidad de Extremadura. Cáceres, España : s.n.
- [REDS] Red Social. Wikipedia. [En línea] http://es.wikipedia.org/wiki/Red_social.
- [LSNE] List of social networking websites. Wikipedia. [En línea] http://en.wikipedia.org/wiki/List_of_social_networking_websites.
- [LINK] LinkedIn. LinkedIn. [En línea] <http://www.linkedin.com/>.
- [JJ09] Jovanovic, Jelena, y otros. The Social Semantic Web in Intelligent Learning Environments: state of the art and future challenges. s.l. : Faculty of Organizational Sciences, University of Belgrade, Belgrade,

- Serbia, 2009. 10.1080/10494820903195140.□
- [UB08] Uldis Bojars, John G. Breslin, Vassilios Peristeras, Giovanni Tummarello, Stefan Decker. Interlinking the Social Web with Semantics. Galway : DERI, National University of Ireland, 2008. 1541-1672/08.
- [CONF] Confianza. Real Academia Española. [En línea] http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=confianza.
- [GAMD] Gambetta, D. Confianza. Wikipedia. [En línea] Can We Trust Trust?. <http://es.wikipedia.org/wiki/Confianza>.
- [CL99] Cornu, Laurence. La confianza en las relaciones pedagógicas. [ed.] Ediciones Novedades Educativas. Buenos Aires : s.n., 1999. págs. p. 19-26. Vol. Construyendo un saber sobre el interior de la escuela.
- [GE98] Gerck, E. Toward Real-World Models of Trust: Reliance on Received Information. The Meta-Certificate Workgroup. [En línea] 1998. <http://mcwg.org/mcg-mirror/trustdef.htm>.
- [PS01] Palmer, Sean B. The Semantic Web: An Introduction. infomesh.net. [En línea] 2001. <http://infomesh.net/2001/swintro/#trustAndProof>.
- [NH93] NIX v. HEDDEN, 149 U.S. 304 (1893). Findlaw. [En línea] <http://caselaw.lp.findlaw.com/scripts/getcase.pl?court=US&vol=149&invol=304>.
- [RAEE] Evaluar. REAL ACADEMIA ESPAÑOLA. [En línea] http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=evaluar.
- [EVAW] Evaluation. Wikipedia. [En línea] <http://en.wikipedia.org/wiki/Evaluation>.
- [HM09] Hurteau, M., Houle, S., & Mongiat, S. How Legitimate and Justified are Judgments in Program Evaluation. 2009. págs. 307-319.
- [LT09] Linman, Terri. 360-degree Feedback: Weighing the Pros and Cons. San Diego State University. [En línea]
- [CERT] Certification. Wikipedia. [En línea] <http://en.wikipedia.org/wiki/Certification>.
- [RAET] Tiempo. REAL ACADEMIA ESPAÑOLA. [En línea] http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=tiempo.
- [TIME] Time. Wikipedia. [En línea] <http://en.wikipedia.org/wiki/Time>.
- [VL05] Viljanen, Lea. Towards an Ontology of Trust. University of Helsinki, Department of Computer Science. Helsinki : In Proceedings of th2nd International Conference on Trust, Privacy and Security in Digital Business (TrustBus'05), 2005.
- [IS06] PRACTICAL ARCHITECTURAL APPROACH FOR COMPOSING EGOCENTRIC TRUST. University of California, Irvine. [En línea]
- [SK03] The EigenTrust Algorithm for Reputation Management in P2P Networks. Sepandar D. Kamvar, Mario T. Schlosser , Hector Garcia-molina. 2003.
- [PS98] Page, Sergey Brin and Lawrence. The anatomy of a large-scale hypertextual Web search engine. Stanford, CA 94305, USA : Computer Science Department, Stanford University, 1998. 0169-7552.
- [RY02] Ratnakar, Yolanda Gil and Varun. Trusting Information Sources One Citizen at a Time. USC Information Sciences Institute. Marina del Rey, CA 90292 : Proceedings of the First International Semantic Web Conference on The Semantic Web , 2002. págs. 162 - 176 . 3-540-43760-6.
- [JS01] Jordi Sabater, Carles Sierra. Social ReGreT, a reputation model based on social relations. Catalonia, Spain : IIIA - Artificial Intelligence Research Institute, CSIC - Spanish Scientific Research Council, Bellaterra, 2001. págs. 44 - 56 .
- [JP02] Josep M. Pujol, Ramon Sangüesa. Reputation Measures based on Social Networks metrics for Multi Agent Systems . Barcelona : Software Department Universitat Politècnica de Catalunya .
- [JG06] Golbeck, Jennifer. Combining Provenance with Trust in Social Networks for Semantic Web Content Filtering. Chicago, Illinois : Proceedings of the International Provenance and Annotation Workshop, 2006.
- [JF06] Jennifer Golbeck, James Hendler. FilmTrust: Movie recommendations using trust in web-based social networks. s.l. : Proceedings of the IEEE Consumer Communications and Networking Conference, January 2006.
- [MICR] Microformats.org. Microformats. [En línea] <http://microformats.org/>.
- [HRXM] HR-XML Consortium. HR-XML Consortium. [En línea] <http://www.hr-xml.org/hr-xml/wms/hr-xml-1-org/index.php?language=2>.
- [RO07] ResumeRDF Ontology Specification. rdfs.org. [En línea] 2007. <http://rdfs.org/resume-rdf/>.
- [SIOC] sioc-project.org. sioc-project.org. [En línea] Digital Enterprise Research Institute, National University of Ireland, Galway. <http://sioc-project.org/>.
- [CG07] Claudio Gutierrez, Carlos Hurtado, and Alejandro Vaisman. Temporal RDF. Universidad de Chile, Universidad de Buenos

- Aires. s.l. : IEEE Transactions on Knowledge and Data Engineering, 2007. vol. 19, no. 2, pp. 207-218.
- [MNDL] McGuinness, Natalya F. Noy and Deborah L. Ontology Development 101: A Guide to Creating Your First Ontology . Stanford, CA : Stanford University.
- [GK09] Eric Gilbert and Karrie Karahalios, Predicting Tie Strength with social media, University of Illinois at Urbana-Champaign, 2009
- [BL03] Brenda Laurel, Peter Lunenfeld. Design Research: Methods and Perspectives. s.l. : The MIT Press, 2003. 0262122634.