

Estado: El preprint no ha sido enviado para publicación

Comparación con grafos de programas de pregrado de estudio en informática del Perú

Jorge L. Alvarado Revata

<https://doi.org/10.1590/SciELOPreprints.1992>

Este preprint fue presentado bajo las siguientes condiciones:

- Los autores declaran que son conscientes de que son los únicos responsables del contenido del preprint y que el depósito en SciELO Preprints no significa ningún compromiso por parte de SciELO, excepto su preservación y difusión.
- Los autores declaran que la investigación que dio origen al manuscrito siguió buenas prácticas éticas y que las aprobaciones necesarias de los comités de ética de investigación se describen en el manuscrito, cuando corresponda.
- Los autores declaran que se obtuvieron los términos necesarios del consentimiento libre e informado de los participantes o pacientes en la investigación y se describen en el manuscrito, cuando corresponde.
- Los autores declaran que la preparación del manuscrito siguió las normas éticas de comunicación científica.
- Los autores declaran que el manuscrito no fue depositado y/o previamente puesto a disposición en otro servidor de preprints o publicado en una revista.
- El autor que presenta declara que todos los autores responsables de la preparación del manuscrito están de acuerdo con este depósito.
- Los autores declaran que si el manuscrito se publicará en el servidor SciELO Preprints, estará disponible bajo licencia [Creative Commons CC-BY](#).
- El manuscrito depositado está en formato PDF.
- El autor que hace el envío declara que las contribuciones de todos los autores están incluidas en el manuscrito.
- Si el manuscrito está siendo evaluado o siendo preparando para su publicación pero aún no ha sido publicado por una revista, los autores declaran que han recibido autorización de la revista para hacer este depósito.

Enviado en (AAAA-MM-DD): 2021-03-13

Postado en (AAAA-MM-DD): 2021-03-22

Comparison with graphs of study plans in computer science from Peru

Jorge Alvarado *

j.revatta@gmail.com

Universidad Nacional Mayor de San Marcos

Lima, Lima, Perú

ABSTRACT

Computer science education in Peru receives different names and different orientations according to the time and place of origin of the profession. Regardless of that, each proposal has curricula, areas of knowledge, skills and professional profiles that establish a varied offer adapted to the country's economic system.

The present study proposes a quantitative approach to identify similarities and differences between curricula of this educational offer. The project incorporates modeling in graphs to convert study plans and international guides into models that allow us to study them in this context. Once a plan or international reference is modeled as a graph, we can study the relationships and thus have a level of approximation between the models. In this way, we can have measures of closeness or distance with the referents or even between the study plans.


CCS CONCEPTS

• **Social and professional topics** → **Computing education programs**;

KEYWORDS

standard guidelines, innovation in computing education, graph theory, maximum common subgraph

ACM Reference Format:

Jorge Alvarado . 2021. Comparison with graphs of study plans in computer science from Peru. In *Proceedings of ACM Conference (Conference'21)*. ACM, New York, NY, USA, 10 pages.

1 INTRODUCCIÓN

El presente estudio propone la aplicación de la teoría de grafos para analizar relaciones entre planes de estudios de programas de informática universitaria en el Perú junto a modelos de guías internacionales de enseñanza de la informática, el objetivo es modelar, visualizar y encontrar información cuantitativa que nos permita posteriormente realizar clasificaciones o tener puntos de partida para procesos de mejora.

Para alcanzar nuestro objetivo usaremos la información pública de planes de estudio de programas de pre-grado, así como información pública de los referentes o guías internacionales de educación

*Facultad de Ingeniería de Sistemas e Informática

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

Conference'21, Marzo 2021, Lima, Perú

© 2021 (), unpublished under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.

en computación. El estudio busca identificar relaciones entre los planes de estudio y los referentes internacionales independiente de la denominación o enfoque que tenga la escuela.

Este estudio expone un proceso y unas técnicas de análisis a nivel micro usando como apoyo la teoría de redes en grafos de manera que cualquier interesado en el diseño de los planes de estudio pueda evaluar medidas de acercamiento a los planes u otras carreras.

2 IDENTIFICACIÓN DEL PROBLEMA

Partiremos planteando tres preguntas principales:

- (1) P1.(H1) ¿Que relaciones podemos identificar en los planes de estudio de programas de informática con otros planes o guías internacionales?
- (2) P2. (H2)¿Que relaciones existe entre los programas de estudio y referentes internacionales en distintos periodos de tiempo?
- (3) P3. (H3)¿Que relaciones existe entre los planes de estudio en distintos periodos de tiempo o distintas casas de estudio?

Como contribución del proceso de investigación se propone un prototipo software para realizar comparaciones y encontrar indicadores cuantitativos, modelando planes de estudios de algunas universidades, así como referentes internacionales. La idea es que en base a este desarrollo se pueda construir una herramienta de comparación para incorporar distintos planes de estudio. El proyecto espera contribuir con nuevos elementos para un mejor entendimiento entre las corrientes formativas de la comunidad informática peruana.

3 LA ADOPCIÓN DE REFERENTES

La adopción de referentes internacionales en las propuestas educativas peruanas requiere de un proceso que permita en primera instancia identificar semejanzas y diferencias para con ello tener una agenda de mejora continua. En las últimas décadas los esfuerzos se han centrado en un enfoque de competencias más que la adopción de áreas de conocimiento. Los esfuerzos para estudiar la adopción de estos referentes se encuentra en los proyectos de [13] donde se busca cierta adaptación de las competencias ACM/IEEE a la región, aunque en ese estudio se presenta una traducción de las competencias de las guías ACM 2005 [15], no se toma en cuenta en realidad el contexto sino una valoración particular por parte de cada institución al cumplimiento o no de las competencias indicadas. Otro enfoque más consensado se encuentra en [5] donde se busca identificar por expertos regionales las competencias genéricas y específicas de profesionales de informática, es decir primero identificar las competencias de cada contexto y luego identificar cierta afinidad y generalidad.

El enfoque propuesto en esta investigación no es estudiar en primera instancia el enfoque de competencias sino unidades de área

de conocimiento que como primer ejemplo serán cursos de planes de estudio así como planes derivados de las guías internacionales. De manera que nos servirán como unidades de trabajo para encontrar equivalencias una vez que se modelen como grafos.

En la presente investigación, usaremos como modelo de guías internacionales a los referentes ACM68 [3], ACM2001 [16], ACM2013 [7], a partir de estos usaremos los cursos comunes y se modelaran como grafos; de igual manera tomaremos planes de estudio de universidades y se modelaran como grafos. Eso nos permitirá utilizar teoría de grafos para encontrar semejanzas o relaciones.

3.1 Las relaciones entre contenidos y áreas de conocimiento

En la investigación de [12] el objetivo es encontrar similitudes entre curriculas de informática para educación escolar. Los autores encuentran valorable el modelamiento con grafos de las curriculas. Para la presente investigación, utilizaremos ese enfoque para modelar las adopciones y cambios en las curriculas de la oferta educativa peruana en informática pero centrados en carreras de pre-grado.

3.2 Modelo basado en grafos

En [14] se define un grafo como:

un conjunto de vértices y una colección de enlaces donde cada enlace conecta un par de vértices.

Se puede representar las propuestas educativas o planes de estudio así como los referentes o guías estándar en un modelamiento con grafos.

En [12] los autores plantean un modelamiento que permita tener en cuenta elementos de conocimiento y competencias vinculados con los estándares. Entre el proceso de modelamiento utilizan texto para hacer anotaciones en las relaciones de los grafos, por ejemplo mencionan identificar vértices y enlaces, donde los enlaces representan relaciones “de requerido por” y “expandido por” para áreas de conocimiento. Los autores introducen el uso de conceptos de la teoría de grafos como métricas para poder realizar comparaciones.

La investigación de [11] modela las curriculas ACM/IEEE como grafos y realiza comparaciones entre los estándares ACM/IEEE 2001, 2008 y 2013, además también hace una comparación con un plan de estudios del mundo real. Los resultados se centran en una métrica basada en comparar el número de nodos y enlaces respecto a cada modelo. El aporte de la presente investigación parte de esa línea de investigación e introduce el uso de un algoritmo de máximo común sub-grafo y de las medidas de distancia basado en isomorfismo para realizar las comparaciones respecto a programas de estudio y los referentes. Es decir introducir una métrica de similitud de grafos, además de aplicar la comparación a los modelos ACM 68 [3], 91 [1]. Por lo que nuestro alcance es mayor en el sentido de intentar responder preguntas de relación en el tiempo con los referentes internacionales; un factor importante para progresar en las confusiones que se derivan de los nombres de las propuestas locales.

En [9] se propone el modelamiento de curriculas como grafos de modo que puede estudiarse métricas como elementos para el análisis. Para el presente análisis se considera relevante identificar un sub-grafo que represente los elementos centrales del área de

informática y buscarlo en los grafos de planes de estudio de las universidades que usaremos en nuestro estudio.

4 SIMILARIDAD DE GRAFOS

En [8], se identifica el marco conceptual de algoritmos de similitud. El teorema para el algoritmo es el siguiente:

Dado dos grafos $G_1(n_1, e_1)$ y $G_2(n_2, e_2)$, con posiblemente diferente número de vértices y aristas, y la red que se forma entre los vértices.

Encontrar, un a) algoritmo que calcule la similitud de los dos grafos y que retorne b) una medida de similitud (un valor real entre 0 y 1).

De igual manera se define el problema de coincidencia de sub-grafos de la siguiente manera:

Dada una serie de tiempo de grafos, donde se encuentran T grafos.

Encontrar a) un algoritmo que busque los sub-grafos aproximados que ocurren en un subconjunto de los T Grafos. b) Donde los grafos aproximados pueden no ocurrir en la mayoría de los puntos de tiempo, en secciones locales de series de tiempo.

4.1 Concepto de isomorfismo y sub-grafos

En [17] se define un isomorfismo de la siguiente manera:

Un sub-grafo G es un grafo cuyos nodos y enlaces se encuentran en G . un grafo G_α es isomorfo a un sub-grafo de un grafo G_β si y solo si existe una correspondencia 1:1 entre el conjunto de nodos del sub-grafo y G_α preserva sus relaciones.

En [4] indica que el isomorfismo de sub-grafos puede ser usado para encontrar si un objeto es parte de otro, o si un objeto está presente en un grupo de objetos. Se menciona el concepto de máximo común sub-grafo, que puede ser usado para medir la similitud de objetos incluso si no existe isomorfismo de grafos o sub-grafos entre los grafos correspondientes. Se busca encontrar el sub-grafo máximo en común entre dos grafos, el mayor representa la similitud.

Una representación teórica de isomorfismo de grafos la encontramos en [6]: Sea $G_1 = (V_1, E_1, \alpha_1, \beta_1)$ y $G_2 = (V_2, E_2, \alpha_2, \beta_2)$ dos grafos. Un isomorfismo de grafos entre G_1 y G_2 es un mapeo biyectivo $f: V_1 \rightarrow V_2$ tal que:

- (1) $\alpha_1(x) = \alpha_2(f(x)), \forall x \in V_1$
- (2) $\beta_1((x, y)) = \beta_2(f(x), f(y)) \forall (x, y) \in E_1$

Si $V_1 = V_2 = \emptyset$, entonces f es llamada un grafo vacío isomorfo.

4.2 Máximo común sub-grafo

De [4] y [6] se define lo siguiente: Sea G_1 y G_2 dos grafos y $G'_1 \subseteq G_1$, $G'_2 \subseteq G_2$. Si existe un grafo isomorfo entre G'_1 y G'_2 , entonces ambos G'_1 y G'_2 son llamados sub-grafo común de G_1 y G_2 . Si no existe otro subgrafo en común de G_1 y G_2 que tenga mas nodos que G'_1 y G'_2 . G'_1 y G'_2 son llamados un MCS de G_1 y G_2 .

Dentro del concepto de común sub-grafo existen variantes como a) Máximo común sub-grafo inducido el cual lo hemos explicado, la búsqueda de la mayor cantidad de vértices, b) En máximo común sub-grafo de enlaces, que permite encontrar el grafo la mayor cantidad de enlaces y que sea isomorfo a ambos grafos comparados.

Como lo indica [4] en usos del mundo real a veces no existe una coincidencia exacta por lo que se considera un grado de error de

tolerancia, usando lo que conocemos como: distancia de edición de grafo. Una operación de edición puede ser una eliminación, inserción o sustitución. Estas operaciones pueden ser aplicadas a nodos o enlaces.

La distancia de edición de dos grafos g y g' es definida como la secuencia mas corta de operaciones de edición que transforman g en g' , la secuencia mas corta indica la mayor similitud entre los grafos. También indican que algunas operaciones de edición pueden tener mas importancia que otras. Por lo que las operaciones de edición tiene un costo asociado individualmente, en nuestro caso no consideraremos un costo asociado para calcular la distancia.

Tanto los conceptos de isomorfismo, de sub-grafos isomorfo y del máximo común sub-grafo son todas instancias de un computo de la edición de distancia de grafos, considerando funciones de costo por operación.

4.3 Métrica de similitud entre grafos

Según [4] define una métrica que no considera las costosas operaciones de edición para convertir un grafo g en g' sino que establece una medida usando el concepto de Máximo común sub-grafo. La medida se establece como la distancia entre dos grafos, usando el número de vértices de un grafo como $|g|$. La formula permite cuantificar la distancia y complementando a 1 se obtiene la similitud entre dos grafos.

$$\delta(g, g') = 1 - \frac{|mcs(g, g')|}{\max(|g|, |g'|)}$$

Este indicador basado en el cálculo del número de nodos del máximo común sub-grafo y el máximo entre grafos comparados nos dará un valor entre 0 y 1 que indica el grado de similitud entre ambos grafos comparados. A menor valor indica una mayor similitud. Es decir si el valor se acerca a 1 es son grafos casi idénticos por el contrario si se acerca a 0 son grafos disimiles.

4.4 Algoritmo de Máximo común sub-grafo

El algoritmo de máximo común sub-grafo es un problema NP por lo que su ejecución se complica conforme crece el número de vértices de los grafos a comparar. El problema es un área de estudio propiamente dicho y de mucha aplicación en el mundo de la química informática. Tiene distintos algoritmos e implementaciones desde la década de los 70 hasta la actualidad. En la presente propuesta usaremos una implementación que identifica el máximo común sub-grafo conectado y usaremos ese valor para el calculo de similitud entre los programas de estudio.

La decisión para elegir el algoritmo que mostraremos se fundamenta en principios pragmáticos que nos permitan alcanzar los resultados experimentales de la solución al problema; no es el propósito de la investigación evaluar el método de implementación del algoritmo de máximo común sub-grafo, dado que el número de nodos a comparar se mantiene en un número bajo que no requerirá un elevado tiempo de espera en la ejecución.

A continuación se explica de modo general una modelo de algoritmo que usamos en el presente proyecto. En nuestra implementación consideramos los sub-grafos conectados pues permite encontrar la estructura de mayor alcance.

Esta implementación no es un algoritmo explicito que corresponda a uno resultante de la literatura por el contrario es un aporte

Algorithm 1: Maximo común subgrafo-conectado

Input: $MCS(G, H)$

Data: G y H dos nodos.

Result: $(MCS - C)$ Maximo común subgrafo-conectado

Data: Grafos G y H

$S = \text{nuevoGrafo}()$ // Crea nuevo grafo S

for $E(n1, n2)$ en G **do**

if H tiene enlaces $E(n1, n2)$ **then**

$S.add(E(n1, n2))$ // Agrega enlace E a S

$\text{Componentes} = \text{ComponentesConectados}(S)$

$\text{ComponenteMasGrande} = \max(\text{Componentes})$

Result: $\text{ComponenteMasGrande}$

pragmático de [10] para encontrar el máximo común sub-grafos conectado.

Los algoritmos de búsqueda de sub-grafos son ampliamente usados en el campo de la informática química o quiminformática, para encontrar patrones químicos en nuestro caso continuamos la línea de encontrar patrones de similitud en los planes de estudio de los programas ofrecidos en algunas universidades del Perú.

4.5 Cuantificación de relaciones y nodos

Un dato importante que también puede considerarse en el objetivo de cuantificar la semejanza o distancia entre las comparaciones, es el uso de una operación simple de conjuntos entre los elementos a comparar. Dado un grafo g y g' calcular cuantos nodos y enlaces son comunes sin que un nodo existente en g esté presente en g' o viceversa. Este simple número puede ayudar a visibilizar los elementos en común entre ambos grafos.

Algorithm 2: Grafo con elementos comunes a G y H

Input: $\text{Comun}(G, H)$

Data: G y H dos nodos.

$T = \text{nuevoGrafo}()$

for $E(x, y)$ en H **do**

if G tiene $E(x, y)$ **then**

$T.add(E(x, y))$

Result: T

Estos valores son importantes para medir la proporción de cursos comunes en ambos grafos de la comparación. También puede servir para medir la proporción de áreas de conocimiento que corresponda en cada curso comparado.

5 HIPOTESIS GENERAL

Como se ha identificado existen dos características importante en la mejora de los planes de estudio, estas son las áreas de conocimientos y el enfoque por competencias. Ambas son necesarias e importantes en la actualidad.

Conference'21, Marzo 2021, Lima, Perú

La presente investigación se concentra en el tema de las áreas de conocimientos, principalmente intentando establecer los cursos de los planes como una primera unidad de comparación para a partir de ello modelar en grafos y establecer criterios de comparación. Se usará una pequeña muestra de datos de planes de estudio al igual que unos modelos de curricula derivados de las guías internacionales. La hipótesis central es poder identificar si existe similitud entre los planes de estudio de distintas escuelas independiente del nombre de la carrera y si se asemejan a planes o guías internacionales independiente de la fecha o etapa de evolución.

La identificación de estas relaciones de áreas de conocimiento requerirá cierto tratamiento de la información de los planes de estudio y de las curriculas internacionales. Se limitará el alcance de la información usada y se aplicará un método de agrupación de los cursos similares o pertenecientes a un mismo cuerpo de conocimiento con un solo identificador o nombre que nos permitirá reconocer como el mismo elemento independiente del plan de estudio o de la guía a comparar.

El aporte de este estudio es el uso del algoritmo Máximo Común Subgrafo así como operaciones de proporcionalidad que nos permita cuantificar y visualizar relaciones o posibles relaciones entre programas de estudio. Esta búsqueda de similitud también es posible realizarla a nivel de cursos como de contenidos, lo que nos permitirá identificar áreas de conocimiento pendientes de abordar o cubiertas por los cursos ofertados.

Tomar a un curso como unidad de comparación puede cuestionarse pues el nombre no es garantía de enseñanza de los mismos objetivos y unidades y sub-unidades de conocimiento. Para el presente estudio nos centraremos solo a criterio experto para encontrar equivalencias en los cursos; en una siguiente investigación con el mismo procedimiento de este trabajo, se puede crear un umbral de similitud comparando áreas, sub-áreas entradas y salidas de un curso.

Nuestra hipótesis general la podemos expresar de la siguiente manera:

Dado un modelo en grafos de un plan de estudios o una guía internacional se identifica una relación entre un plan de estudios G_1 y G_2 como la métrica entre ambos modelos usando el concepto de similaridad grafos. De modo que la similaridad es significativa.

Definiremos un modelamiento siguiendo los siguientes pasos:

- (1) Usaremos el marco teórico de sub-grafos y coincidencia de gráfos para modelar los planes de estudio y/o curriculas (Referentes) para realizar ejercicios de comparación usando como métrica la similaridad entre subgrafos. Visualizaremos también el máximo común subgrafo conectado como un elemento de similitud entre planes y/o curriculas, así como las relaciones en común para dar mayor pistas de las similitudes o diferencias.
- (2) La data será depurada y/o preparada para centrarse en los cursos en contexto informático y de matemáticas, se descartan cursos de gestión, ciencia general y sociales a pesar que son temas que influyen en las competencias modernas, lo que se busca es encontrar la similaridad en áreas centrales de la informática y las propuestas educativas del país.

- (3) Se modela los planes o guías que usaremos para el alcance de la investigación en grafos y se usará un diccionario de nombres que agrupen áreas comunes o cursos similares. Se crea una Tabla de Comparación. Tanto El Plan como el Referente tendrán una tabla de comparación que permita identificarlos usando un nemónico que represente el curso o área de conocimiento que abarca la oferta educativa.
- (4) Aplicar los algoritmos a la codificación de los planes y guías, realizar las comparaciones usando dos modelos: modelo 1) se identifica los cursos de los planes de estudio con equivalencias a ejemplos de curriculas internacionales; en caso de curriculas entre universidades se crea un Id y nombre común a los cursos equivalentes y se modelan los grafos usando ese identificador. En el modelo 1, el mapeo de relaciones se respeta la matriz de relaciones original de la curricula internacional. En el modelo 2 Se compara el plan de estudio local respetando sus relaciones originales. Se presume que bajo esa premisa el subgrafo conectado sea menor o incluso no se pueda identificar.
- (5) Aplicar los algoritmos a la codificación de áreas de conocimiento de los cursos comparando el mapeo respecto a modelos de cursos de las guías internacionales, de manera que un curso de una universidad X cuantifique su aproximación a un mapeo de (*UnidadConocimiento*) de un curso considerado ejemplo de una guía internacional.
- (6) Cuantificar el número de nodos y relaciones por cada subgrafo de comparación y tabular los resultados. Estos datos nos permitirán cuantificar tanto los cambios en las curriculas como las similitudes entre los planes existentes y las propuestas educativas locales. Cuantificar los enlaces comunes entre ambos grafos para visibilizar la relaciones.
- (7) Mostrar la evidencia numérica que apruebe o rechace la hipótesis general y secundarias.

No se ha considerado evaluar o comparar competencias pues no existe información pública de estas en las universidades que hemos tomado como muestras en el presente estudio. En el caso de las áreas de conocimiento que se compararan a nivel de cursos, estos se infieren de revisión de planes de estudio de los cursos que tengan información pública.

6 HIPOTESIS ESPECÍFICAS

A continuación usaremos la hipótesis principal para responder preguntas vinculadas con el tema que afectan el proceso de desarrollo de las profesiones universitarias en informática. Esta pregunta principal la podemos desagregar en las preguntas secundarias que generan las hipótesis específicas.

- (1) P2. (H2) Dado un grafo G_1 que representa una profesión de ingeniería de sistemas u otra propuesta existente en la oferta peruana y G_2 un referente internacional de ACM/IEEE u otro. Se identifica que la similaridad es significativa.
- (2) P3. (H3) Dado un grafo G_1 que representa el plan de estudio de una carrera vinculada a la informática de un año X y G_2 un plan de la misma carrera pero del año Y . Se Identifica que la similaridad es significativa. Una modificación también podría permitir comparar dos carreras locales de distintos años y validar la similaridad.

Dado que usaremos una métrica basada en el Máximo común subgrafo, el concepto de significancia pasa por establecer un valor que indique la presencia o no de similaridad, en este caso tomaremos un valor 0.1 como el valor que indique o no la presencia de significancia de similaridad. Es decir una métrica que supere ese valor indicaría cierto grado de similaridad, uno que sea menor un similaridad no significativa.

Otro elemento que podemos considerar es la proporción en una de los planes o guías respecto a la cantidad de nodos en común. En este caso usaremos una proporción 0.1 para indicar cierto grado de similaridad por nodos.

7 POBLACIÓN DE ESTUDIO

La población de nuestro estudio es un plan de estudios de una carrera universitaria en informática. Los postulantes, estudiantes de primeros años, utilizan la información de los planes para tomar decisiones sobre que carrera elegir o que cursos son importantes dentro de la propuesta. Los planes de estudio cambian en el tiempo y son dinámicos por que reflejan por un lado la interpretación institucional de la propuesta pero por otro lado también son un reflejo del enfoque y de las posibilidades de una institución. Por lo tanto son una fotografía de la propuesta formativa, de ahí que también sea importante identificar si los planes se asemejan o no a un modelo estandar de plan de estudio, estos se consideran los acuerdos internacionales que se han diseñado con criterio experto por instituciones dedicadas a ello.

Las personas que puede tener interés en este estudio va desde académicos que evalúan el progreso o desarrollo de los planes de estudio; estudiantes postulantes a las carreras y que se enfrentan a dilemas de decisión sobre que carrera elegir; a estudiantes de las carreras que llevan matriculados para conocer las similitudes o diferencias que existen respecto a su formación y las impartidas por otros centros de enseñanza. Finalmente puede ser de interés para aquellos que estudian el desarrollo y evolución de la educación en informática del país.

8 TAMAÑO DE LA MUESTRA

En el caso de la presente investigación, usaremos información de las siguientes universidades ubicadas en ciudades de marcada tendencia de cambios en la oferta educativa:

Carrera	Universidad	Año plan
Ingeniería informática	Universidad Católica	2020
Ingeniería de Sistemas	Universidad San Marcos	2009
Ingeniería de Sistemas	Universidad San Marcos	2017

Table 1: Matriz de planes de estudio considerados en el estudio

Perfil	Organización	Graduación	Año guía
Informática	ACM	Pre-grado	1968
Informática	ACM/IEEE	Pre-grado	2003
Informática	ACM/IEEE	Pre-grado	2013

Table 2: Matriz de guías internacionales

9 SELECCIÓN DE LA MUESTRA

Las muestras seleccionadas corresponden a un grupo pequeño de planes de estudio de distintos años de algunas universidades que ofrecen programas de informática, en el caso de las guías internacionales se ha seleccionado planes de ejemplo derivados de los referentes. La idea central es poder modelar estas propuestas para alcanzar una representación de grafos que nos permita obtener información cuantitativa del modelo.

Se ha buscado en la oferta pública de las universidades los documentos que resumen los planes de estudio vigentes a la fecha de la realización de la investigación. Se han descargado los documentos públicos y se almacenan en carpetas independientes por universidad. La selección podría ser aleatoria pero dada la poca disponibilidad de recursos de base de datos estandarizados para los planes se ha optado elegir propuestas de los perfiles comunes en la oferta educativa: Ingeniería informática, ingeniería de sistemas, ciencia de la computación.

De igual manera con respecto a los perfiles internacionales que sirven de referentes para innovar o estandarizar los programas de estudio, se van a considerar a los marcos referenciales indicados en 2

10 TÉCNICA ANALÍTICA

Para el análisis de identificación de relaciones usaremos la teoría de grafos modelando los planes, guías internacionales y cursos siguiendo una representación de grafos. Posteriormente usaremos propiedades de la teoría de grafos para encontrar similaridad, distancias, sub-grafos comunes y valores cuantitativos que indiquen la cantidad de relaciones comunes entre ambos modelos a comparar. Usaremos el modelamiento de grafos y los algoritmos respectivos para identificar similaridad y valore cuantitativos.

Nuestra hipótesis central es la siguiente:

- (1) P1. (H1) Dado un modelo en grafos de un plan de estudios o una guía internacional encontrar la relación entre un plan de estudios G_1 y G_2 como la métrica entre ambos modelos usando el concepto de similaridad grafos. De modo que la similaridad es significativa.

Dada la métrica que consideramos en el estudio:

$$\delta(g, g') = 1 - \frac{|mcs(g, g')|}{\max(|g|, |g'|)}$$

Que a su vez depende de encontrar el Máximo común sub-grafo MCS que en el presente estudio depende de un algoritmo indicado en 4.4. Con los datos indicados previamente modelaremos planes, guías internacionales cursos como grafos y encontraremos la métrica para validar las respectivas hipótesis propuestas.

Las hipótesis secundarias se enumeran a continuación:

- (1) P2. (H2) Dado un grafo G_1 que representa una profesión de ingeniería de sistemas o informática y G_2 un referente internacional de ACM/IEEE u otro. Identificar que la similaridad es significativa.
- (2) P3. (H3) Dado un grafo G_1 que representa el plan de estudio de una carrera vinculada a la informática de un año X y G_2 un plan de la misma carrera pero del año Y. Identificar que la similaridad es significativa.

Para validar las hipotesis tendremos 2 métricas.

Conference'21, Marzo 2021, Lima, Perú

El nivel de similitud entre 2 grafos como el valor $z = 0.1$ como el valor que indique o no la presencia significativa de similitud. Un valor mayor a z refleja un nivel aceptable de similitud, un valor menor indica que no existe similitud.

Otro elemento que podemos considerar es la proporción en una de los planes o guías respecto a la cantidad de nodos en común. En este caso usaremos una proporción $t = 0.1$ para indicar cierto grado de similitud por nodos. Esto significa que un valor menor a t significa que no existe similitud en cantidad de nodos en común.

11 PRUEBAS DE HIPÓTESIS

Tal como planteamos en el capítulo anterior de metodología, se sigue una serie de pasos para codificar los datos en grafos. En la siguiente tabla se muestra los datos recolectados de los planes y guías consideradas para el experimento.

Caso	Modelo	G(acm68)-n	G(acm68)-e	H(sis2009)-n	H(sis2009)-e
1	model1	30	39	41	34
1	model2	30	39	42	32
Caso 2	Modelo model1	G(acm68)-n 30	G(acm68)-e 39	H(sis2017)-n 38	H(sis2017)-e 25
2	model2	30	39	38	24
Caso 3	Modelo model1	G(sis2009)-n 30	G(sis2009)-e 39	H(sis2017)-n 38	H(sis2017)-e 25
3	model2	30	39	38	24
Caso 4	Modelo model1	G(acm2013)-n 30	G(acm2013)-e 39	H(sis2017)-n 38	H(sis2017)-e 25
4	model2	30	39	38	24
Caso 5	Modelo model1	G(acm2013)-n 26	G(acm2013)-e 20	H(pucp2020)-n 60	H(pucp2020)-e 20
5	model2	26	20	60	17

Table 3: Tabulación de datos sobre comparaciones

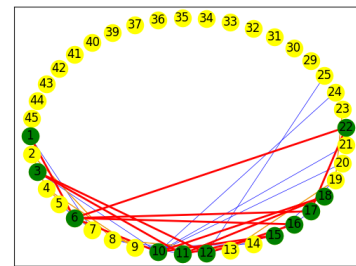


Figure 1: Sis2009 ACM68 MCS modelo 1

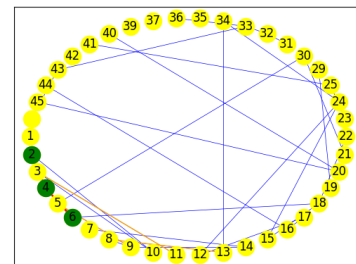


Figure 2: Sis2009 ACM68 MCS modelo 2

Caso	Modelo	G(acm68)-n	G(acm68)-e	H(sis2009)-n	H(sis2009)-e
1	model1	30	39	41	34
1	model2	30	39	42	32

Table 4: Tabulación de datos sobre comparaciones

12 PRESENTACIÓN DE RESULTADOS

Presentaremos ahora los resultados organizados por casos. Es importante explicar algunas características de las visualizaciones:

- (1) Por caso se dan dos comparaciones, denominadas *Modelo₁* y *Modelo₂*.
- (2) Los enlaces comunes a ambos grafos aparecen en color naranja.
- (3) Los enlaces aparecen en color azul.
- (4) Los nodos del modelo en amarillo.
- (5) Los nodos del MCS en color verde

12.1 Caso 1 Hipotesis P2

A continuación se muestran las representaciones gráficas y los resultados del *Caso₁* de la comparación entre un plan de estudio: Sistemas 2009 y la guía internacional ACM 68.

Caso	Modelo	MCSsubgrado	similitud	nodos	enlaces
1	<i>modelo₁</i>	0.64285714	$z = 0.35714286$	11	13
1	<i>modelo₂</i>	0.9285714286	$z = 0.07142857143$	3	2

Table 5: Comparación G-ACM 68 y H-Sistemas 2009

Caso	Modelo	nodosCom	enlacesCom	proporciónG	proporciónH
1	<i>modelo₁</i>	20	21	$t_1 = 0.6666666667$	$t_2 = 0.487804878$
1	<i>modelo₂</i>	9	5	$t_1 = 0.3$	$t_2 = 0.2142857143$

Table 6: Tabla de resultados de proporcionalidad

Hipotesis	Caso	modelo	Similitud	ProporciónG	ProporciónH
P2	1	modelo1	$z > 0.1$	$t_1 > 0.1$	$t_2 > 0.1$
P2	1	modelo2	$z < 0.1$	$t_1 > 0.1$	$t_2 < 0.1$

Table 7: Tabla de Hipotesis vs Resultados

12.2 Caso 2 Hipotesis P2

A continuación se muestran las representaciones gráficas y los resultados del *Caso₂* de la comparación entre un plan de estudio: Sistemas 2017 y la guía internacional ACM 68.

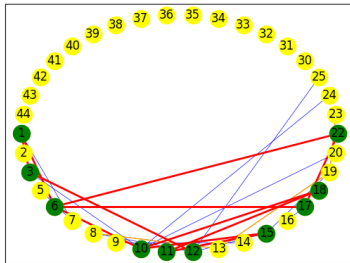


Figure 3: Sis2017 ACM68 MCS modelo 1

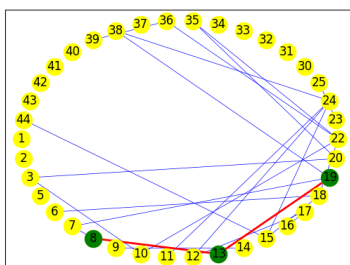


Figure 4: Sis2017 ACM68 MCS modelo 2

Caso	Modelo	G(acm68)-n	G(acm68)-e	H(sis2009)-n	H(sis2009)-e
2	modelo1	30	39	38	25
2	modelo2	30	39	38	24

Table 8: Tabulación de datos sobre comparaciones

Caso	Modelo	MCSsubgrado	similaridad	nodos	enlaces
2	modelo ₁	0.736842105263158	$z = 0.2631578947$	10	11
2	modelo ₂	0.9210526316	$z = 0.07894736842$	3	2

Table 9: Comparación G-ACM 68 y H-Sistemas 2009

Caso	Modelo	nodosCom	enlacesCom	proporcionG	proporcionH
2	modelo ₁	14	14	$t_1 = 0.4666666667$	$t_2 = 0.3684210526$
2	modelo ₂	3	2	$t_1 = 0.1$	$t_2 = 0.07894736842$

Table 10: Tabla de resultados de proporcionalidad

Hipotesis	Caso	modelo	Similaridad	ProporcionG	ProporcionH
P2	2	modelo1	$z > 0.1$	$t_1 > 0.1$	$t_2 > 0.1$
P2	2	modelo1	$z < 0.1$	$t_1 > 0.1$	$t_2 > 0.1$

Table 11: Tabla de Hipotesis vs Resultados

12.3 Caso 3 Hipotesis P3

A continuación se muestran las representaciones gráficas y los resultados del *Caso₃* de la comparación entre un plan de estudio: Ingeniería de Sistemas 2009 e Ingeniería de Sistemas 2017.

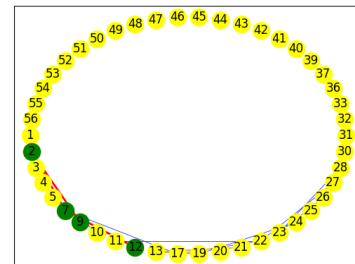


Figure 5: Sis2009 Sis2017 MCS modelo 1

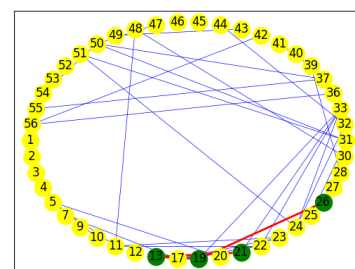


Figure 6: Sis2009 Sis2017 MCS modelo 2

Caso	Modelo	G(Sis2009)-n	G(Sis2009)-e	H(Sis2017)-n	H(Sis2017)-e
3	modelo1	30	39	38	25
3	modelo2	30	39	38	24

Table 12: Tabulación de datos sobre comparaciones

Caso	Modelo	MCSsubgrado	similaridad	nodos	enlaces
3	modelo ₁	0.914893617	$z = 0.08510638298$	4	3
3	modelo ₂	0.914893617	$z = 0.08510638298$	4	3

Caso	Modelo	nodosCom	enlacesCom	proporcionG	proporcionH
3	modelo ₁	7	5	$t_1 = 0.1489361702$	$t_2 = 0.152173913$
3	modelo ₂	4	3	$t_1 = 0.08510638298$	$t_2 = 0.08695652174$

Table 13: Tabla de resultados de proporcionalidad

Conference'21, Marzo 2021, Lima, Perú

Hipotesis	Caso	modelo	Similaridad	ProporcionG	ProporcionH
P3	3	modelo1	$z < 0.1$	$t_1 > 0.1$	$t_2 > 0.1$
P3	3	modelo1	$z < 0.1$	$t_1 < 0.1$	$t_2 < 0.1$

Table 14: Tabla de Hipotesis vs Resultados

Hipotesis	Caso	modelo	Similaridad	ProporcionG	ProporcionH
P2	4	modelo1	$z > 0.01$	$t_1 > 0.1$	$t_2 > 0.1$
P2	4	modelo1	$z > 0.01$	$t_1 > 0.1$	$t_2 > 0.1$

Table 17: Tabla de Hipotesis vs Resultados

12.4 Caso 4 Hipotesis P2

A continuación se muestran las representaciones gráficas y los resultados del *Caso4* de la comparación entre un plan de estudio: Ingeniería de sistemas 2017 y ACM 2013.

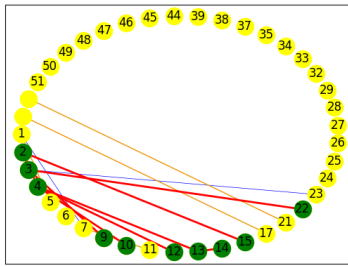


Figure 7: Sis2017 y ACM 2013 MCS modelo 1

12.5 Caso 5 Hipotesis P2

A continuación se muestran las representaciones gráficas y los resultados del *Caso5* de la comparación entre un plan de estudio de ingeniería informática y una guía internacional ACM 2013.

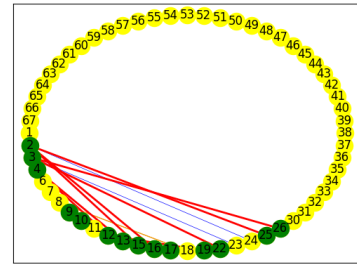


Figure 9: IngInform2020 y ACM 2013 MCS modelo 1

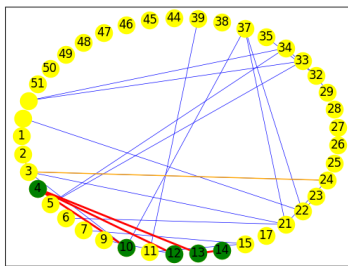


Figure 8: Sis2017 y ACM 2013 MCS modelo 2

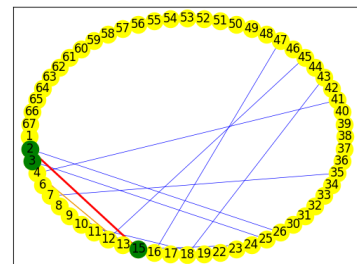


Figure 10: IngInform2020 y ACM 2013 MCS modelo 2

Caso	Modelo	G(Sis2017)-n	G(Sis2017)-e	H(ACM2013)-n	H(ACM2013)-e
4	model1	30	39	38	25
4	model2	30	39	38	24

Table 15: Tabulación de datos sobre comparaciones

Caso	Modelo	G(IngInform2020)-n	G(IngInform2020)-e	H(ACM2013)-n	H(ACM2013)-e
5	model1	26	20	60	20
5	model2	26	20	60	17

Table 18: Tabulación de datos sobre comparaciones

Caso	Modelo	MCSsubgrado	similaridad	nodos	enlaces
4	modelo ₁	0.756097561	$z = 0.243902439$	10	9
4	modelo ₂	0.8780487805	$z = 0.1219512195$	5	4

Table 16: Tabla de resultados de proporcionalidad

Caso	Modelo	MCSsubgrado	similaridad	nodos	enlaces
5	modelo ₁	0.7666666666666666	$z = 0.2333333333$	18	16
5	modelo ₂	0.95	$z = 0.05$	5	3

Table 19: Tabla de resultados de proporcionalidad

Caso	Modelo	nodosCom	enlacesCom	proporcionG	proporcionH
4	Modelo ₁	18	13	$t_1 = 0.75$	$t_2 = 0.4390243902$
4	Modelo ₂	8	6	$t_1 = 0.3333333333$	$t_2 = 0.1951219512$

Caso	Modelo	nodosCom	enlacesCom	proporcionG	proporcionH
5	Modelo ₁	18	16	$t_1 = 0.6923076923$	$t_2 = 0.3$
5	Modelo ₂	5	3	$t_1 = 0.1923076923$	$t_2 = 0.08333333333$

Conference'21, Marzo 2021, Lima, Perú

de comparación con un nivel de probabilidad para realizar las comparaciones. Otra oportunidad que puede derivarse de este estudio es la inclusión de programas de estudio de un número mayor de universidades así como realizar comparaciones entre estos planes de estudio.

REFERENCES

- [1] 1991. *Computing Curricula 1991: Report of the ACM/IEEE-CS Joint Curriculum Task Force*. Technical Report. New York, NY, USA.
- [2] Jorge Alvarado. 2021. *Heramienta de comparación @ONLINE*. <https://comparagraphfront.herokuapp.com>
- [3] William F. Atchison, Samuel D. Conte, John W. Hamblen, Thomas E. Hull, Thomas A. Keenan, William B. Kehl, Edward J. McCluskey, Silvio O. Navarro, Werner C. Rheinboldt, Earl J. Schweppe, William Viavant, and David M. Young. 1968. Curriculum 68: Recommendations for Academic Programs in Computer Science: A Report of the ACM Curriculum Committee on Computer Science. *Commun. ACM* 11, 3 (mar 1968), 151–197. <https://doi.org/10.1145/362929.362976>
- [4] Horst Bunke. 2000. Graph matching: Theoretical foundations, algorithms, and applications. In *In Proceedings of Vision Interface 2000, Montreal*. 82–88.
- [5] José Lino Contreras Véliz, Javier Alanoca Gutiérrez, Jamil Salem Barbar, Jorge Enrique Quevedo Reyes, Gabriela Garita, Roberto Sepúlveda Lima, Cecilia Hinojosa Raza, Héctor José Duarte Pavón, Alma Patricia Chávez Cervantes, Augusto Enrique Estrada Quintero, Diana Bernal, María Elena García, José Antonio Pow-Sang, and Laura González. 2013. *Tuning LA: Educación Superior en América Latina: reflexiones y perspectivas en Informática*. Technical Report. http://tuningacademy.org/wp-content/uploads/2014/02/RefInformatics_LA_SP.pdf
- [6] Xinbo Gao, Bing Xiao, Dacheng Tao, and Xuelong Li. 2010. A survey of graph edit distance. *Pattern Analysis and Applications* 13, 1 (01 Feb 2010), 113–129. <https://doi.org/10.1007/s10044-008-0141-y>
- [7] Association for Computing Machinery (ACM) Joint Task Force on Computing Curricula and IEEE Computer Society. 2013. *Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science*. Association for Computing Machinery, New York, NY, USA.
- [8] Danai Koutra, Aaditya Ramdas, Ankur Parikh, and Jing Xiang. 2011. Algorithms for Graph Similarity and Subgraph Matching.
- [9] Jay M. Lightfoot. 2010. A Graph-Theoretic Approach to Improved Curriculum Structure and Assessment Placement. In *Communications of the IIMA*. <https://scholarworks.lib.csusb.edu/ciima/vol10/iss2/5>
- [10] Magin. 2019. <https://stackoverflow.com/questions/43108481/maximum-common-subgraph-in-a-directed-graph>
- [11] Linda Marshall. 2012. A Comparison of the Core Aspects of the ACM/IEEE Computer Science Curriculum 2013 Strawman Report with the Specified Core of CC2001 and CS2008 Review. In *Proceedings of Second Computer Science Education Research Conference (Wroclaw, Poland) (CSERC '12)*. Association for Computing Machinery, New York, NY, USA, 29–34. <https://doi.org/10.1145/2421277.2421281>
- [12] S. Pasterk and A. Bollin. 2017. Graph-based analysis of computer science curricula for primary education. In *2017 IEEE Frontiers in Education Conference (FIE)*. 1–9. <https://doi.org/10.1109/FIE.2017.8190610>
- [13] Teófilo J. Ramos, Osvaldo M. Micheloud, Richard Painter, and Moshe Kam. 2013. Nomenclatura recomendada por el IEEE para programas Universitarios del área computacional en latinoamerica. (2013). <http://www.accreditation.org/sites/default/files/Latin-American-Computing-Nomenclature-Documents-English.pdf>
- [14] Robert Sedgewick and Kevin Wayne. 2011. *Algorithms* (4th ed.). Addison-Wesley Professional.
- [15] Russell Shackelford, Andrew McGettrick, Robert Sloan, Heikki Topi, Gordon Davies, Reza Kamali, James Cross, John Impagliazzo, Richard LeBlanc, and Barry Lunt. 2006. Computing Curricula 2005: The Overview Report. *SIGCSE Bull.* 38, 1 (March 2006), 456–457. <https://doi.org/10.1145/1124706.1121482>
- [16] CORPORATE The Joint Task Force on Computing Curricula. 2001. Computing Curricula 2001. *J. Educ. Resour. Comput.* 1, 3es (Sept. 2001), 1–es. <https://doi.org/10.1145/384274.384275>
- [17] J. R. Ullmann. 1976. An algorithm for subgraph isomorphism. *JOURNAL OF THE ACM* 28, 1 (1976), 31–42.

A DECLARACIÓN DE CONFLICTO DE INTERESES

el autor declara no tener conflictos de intereses por declarar. El artículo no tiene ningún interés financiero por reportar. Certifico que el trabajo es un proyecto de investigación original y que no se encuentra en revisión en alguna publicación.

Firma: Jorge Alvarado. 15/03/2021.