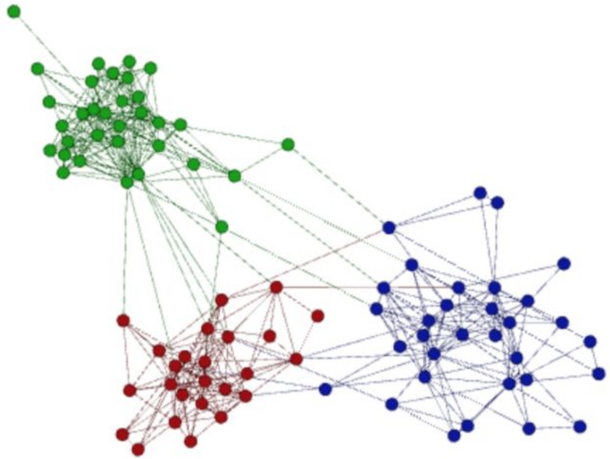
Redes de proyecto por camino crítico
Clase 11

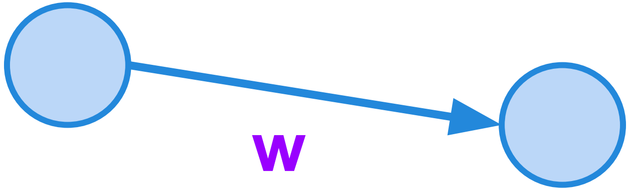
Investigación Operativa UTN FRBA 2020
Curso: I4051
Equipo: Juan Piro, Milagros Bochor, Gabriel Boso, Rodrigo Maranzana
Docente: Martín Palazzo

Agenda clase 11

- Gafos: aplicaciones
- PERT
- Ejercicios PERT



Grafo de un proyecto



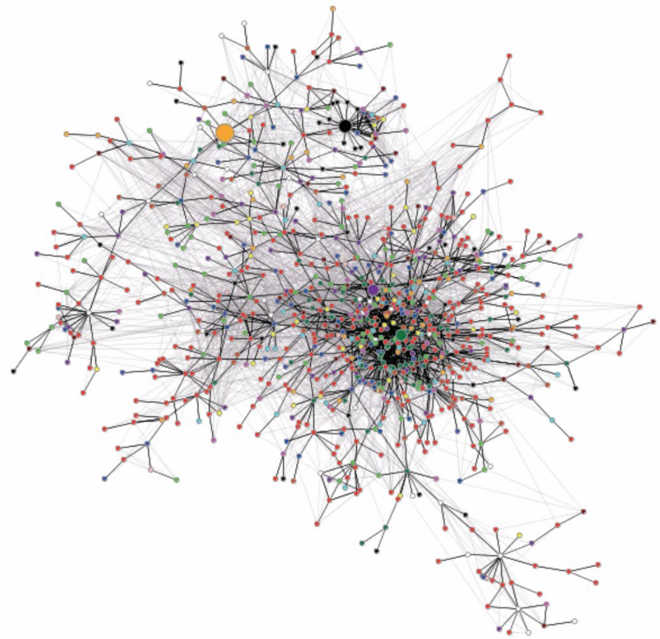
En un contexto de proyectos cada arco representa una tarea. El **peso** (weight) de cada arco puede estar caracterizado visualmente con su grosor (y no por su longitud). En la matriz de adyacencia los pesos de los arcos estan representados en los elementos de la matriz. Los pesos pueden representar:

- Tiempo
- Distancia
- Flujo
- Resistencia
- Costo
- Interacciones

Aplicaciones de Grafos: redes sociales

Un caso de aplicación de grafos es el análisis de redes sociales (digitales, ya que existen las redes sociales fisicas). ¿Qué podrían significar los arcos en este grafo? ¿y que representan sus pesos?

Aplicaciones de Grafos: familiar strangers



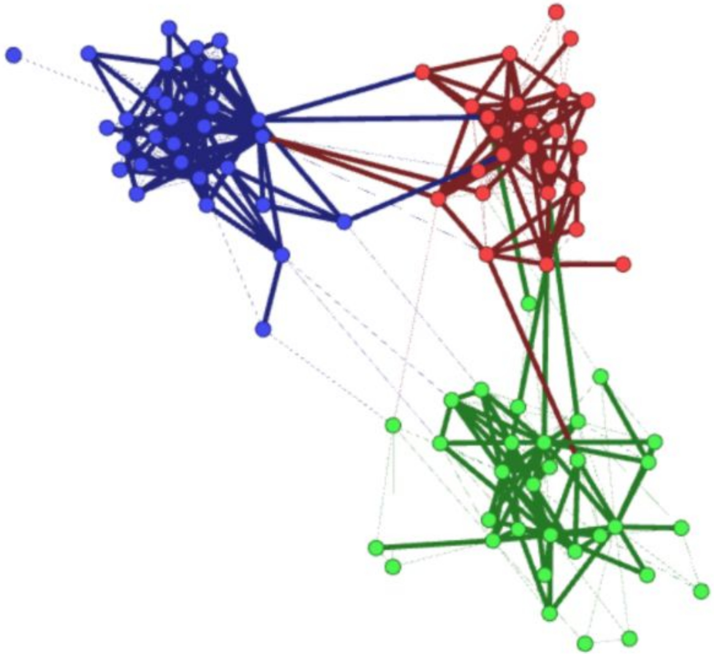
*A Social Identity Approach to Identify Familiar Strangers in a Social Network

Otro caso de teoría de grafos es para la detección de 'Familiar Strangers' (extraños familiares) y se determinan como aquellas personas que no conoces pero tenes algo en común (existe algún grado de similaridad).

En el caso de familiar strangers podemos decir que dos personas comparten dicha categoría si por ejemplo comparten espacio y tiempo o solo alguna de ellas. Ejemplo: si todos los dias tomas el mismo colectivo a la misma hora posiblemente veas las mismas caras aunque no tengas un vinculo directo con esas personas. Esto genera que si te cruzas con esas personas en un entorno totalmente distinto posiblemente tengas mas probabilidad de contacto (ey! Te conozco del bondi).

En este tipo de situaciones un arco entre dos individuos es realizado cuando ambos comparten la misma ventana de espacio y tiempo. El peso del arco dependerá de otros factores como la cantidad de encuentros.

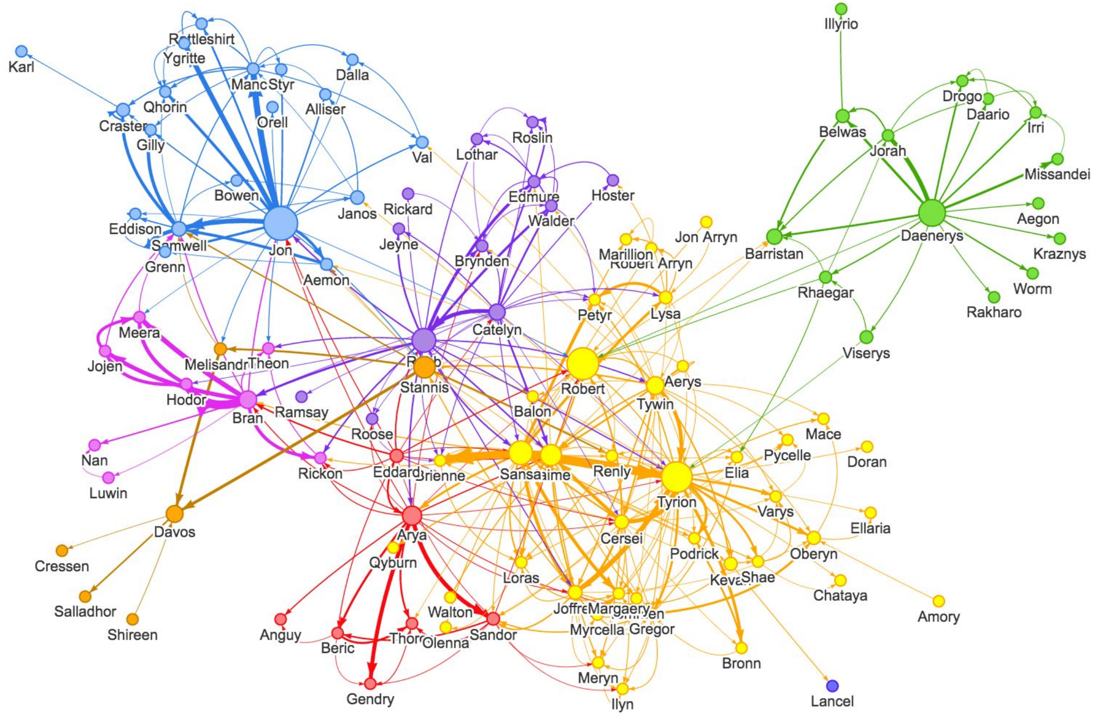
Aplicaciones de Grafos: community detection



*Spectral Clustering in Dynamic Graphs, Martin Palazzo

Existen algoritmos que permiten encontrar sub-estructuras mas compactas con mayor densidad de conexiones dentro de un grafo. A estas sub-estructuras las llamaremos "*clusters*" o "*comunidades*". Encontrar comunidades tiene múltiples aplicaciones como segmentación de clientes en un contexto de marketing, poder encontrar subtipos de una misma patología (por ejemplo existen 3 subtipos de cancer de mamas) y poder entender cuan facil por ejemplo puede viajar un mensaje (o virus) de una comunidad a otra.

Aplicaciones de Grafos: community detection



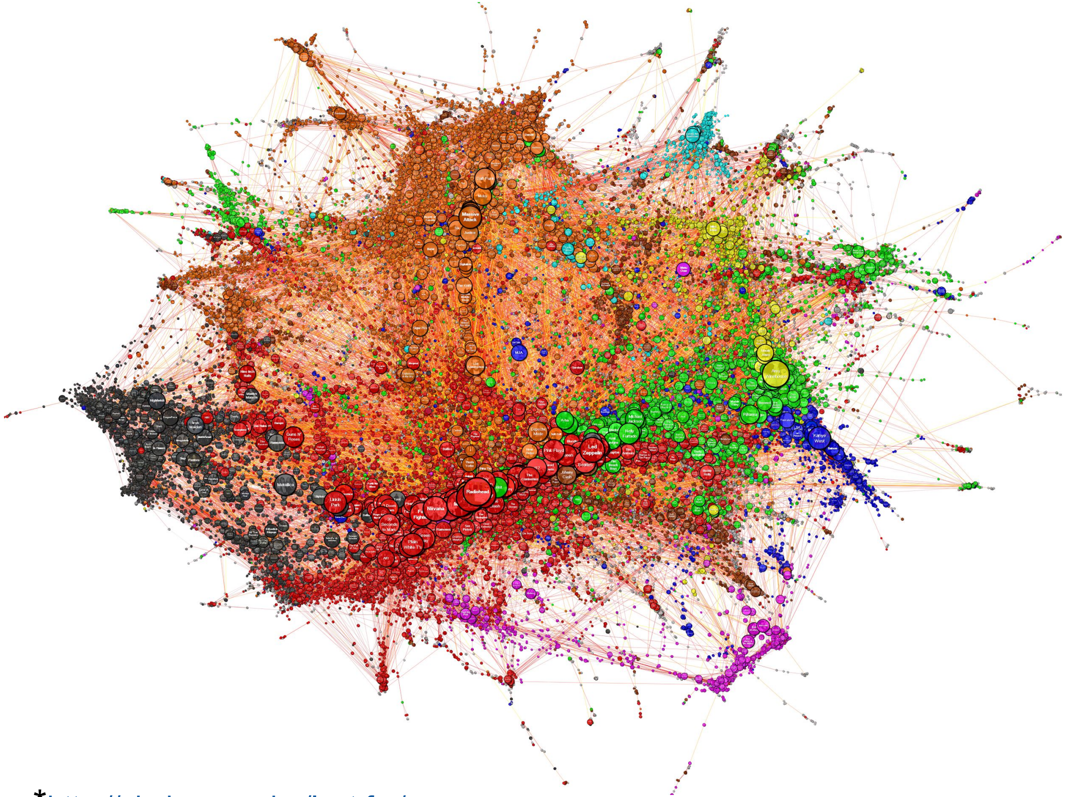
Por ejemplo, si construimos un grafo con los personajes de Games of Thrones podemos encontrar comunidades de personajes!!

En este grafo cada nodo representa un personaje y cada arco indica si interactuaron o no. De esta manera se pueden observar 8 comunidades de personajes. El peso de los arcos indica la cantidad de interacciones que tuvieron esos personajes.

Esto muestra una vez mas como los grafos pueden usarse para modelar las dinámicas humanas.

*https://neo4j.com/blog/graph-of-thrones/

Aplicaciones de Grafos: community detection



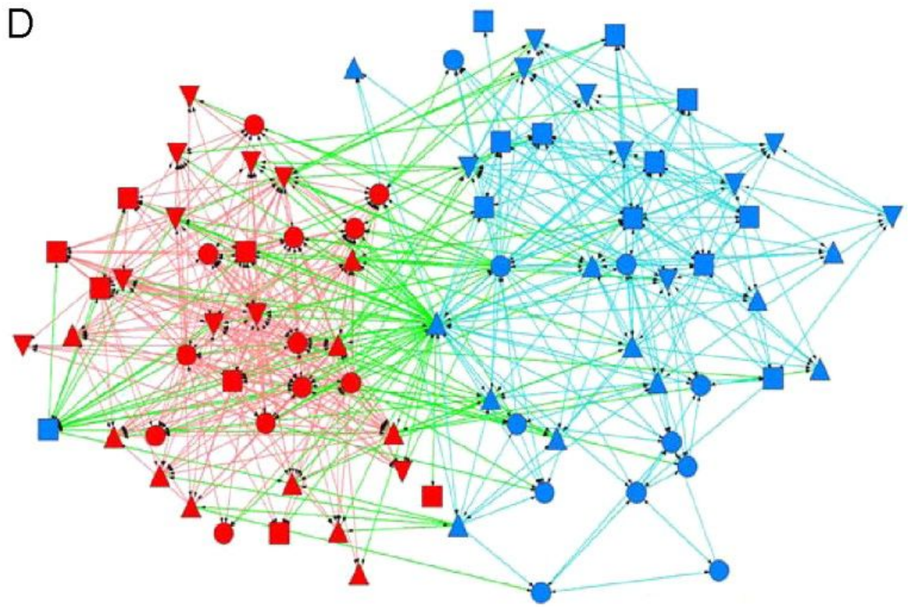
En este caso se muestra un grafo con las bandas del sitio last.fm. En este grafo cada nodo es un artista y los arcos representan la similaridad entre artistas.

La similaridad es calculada en función de características propias del artista (estilo) y en las preferencias entre los usuario. Por ejemplo, muchos usuarios de Tame Impala (Psych-Rock Progresivo contemporáneo) pueden escuchar The Chemical Brothers (Electronica).

El grafo muestra claramente mas de comunidades, algunas mas grandes que otras.

*http://sixdegrees.hu/last.fm/

Aplicaciones de Grafos: spread disease



*Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza (https://www.pnas.org/content/108/7/2825)

Los grafos también pueden utilizarse para modelar la transmisión de enfermedades en un tejido social.

El grafo visualizado corresponde a la dinámica social de una escuela en el año 2009 donde se originó uno de los brotes de la pandemia H1N1 influenza.

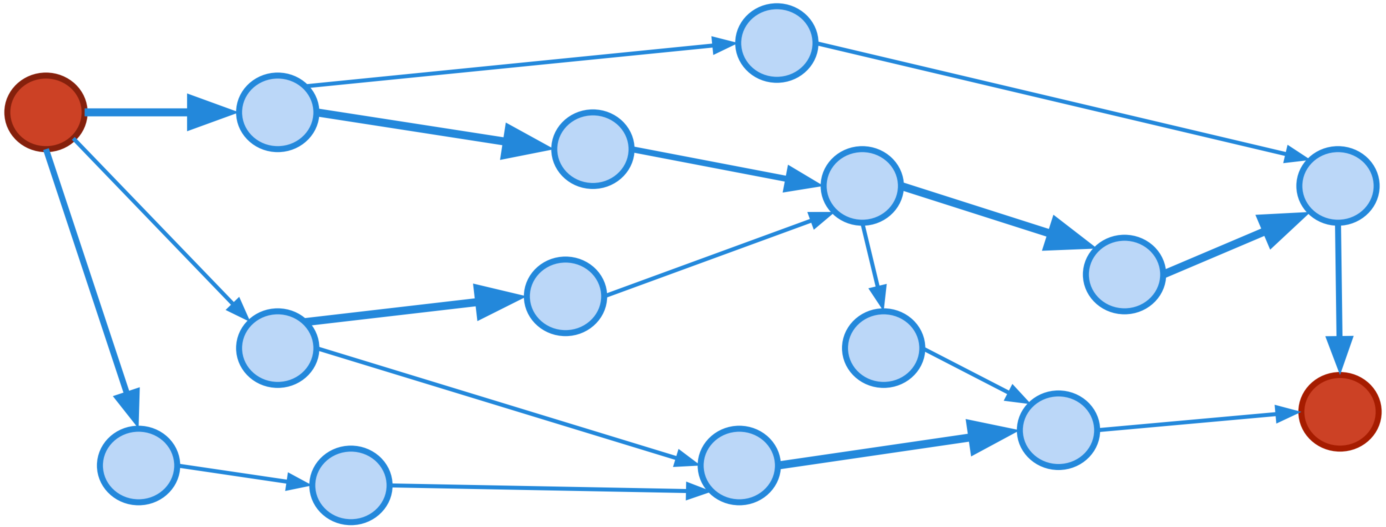
Entender la conectividad de la red ayuda a entender la velocidad de propagación de la patología a través de toda la red.

De igual manera uno podría realizar el seguimiento de un mensaje, un rumor, a través de una red social.

Librerías interesantes de grafos en Python

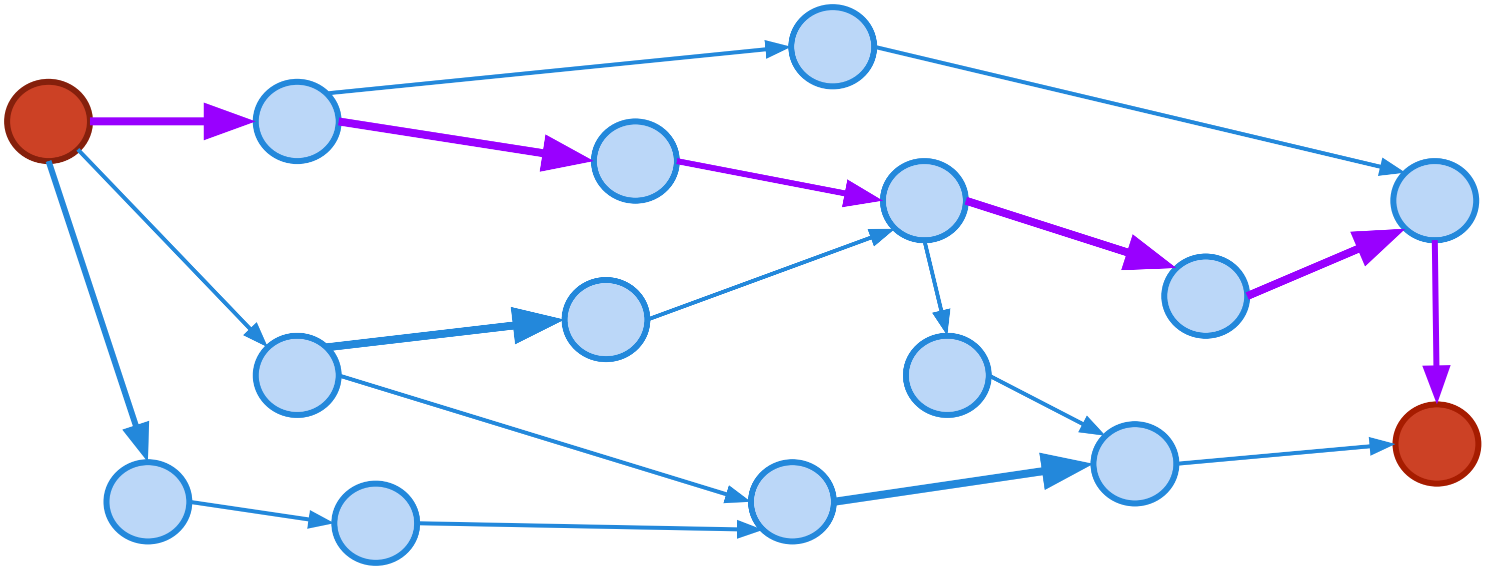
- NetworkX https://networkx.github.io/
- Littlebaloffur https://github.com/benedekrozemberczki/littleballoffur
- Gephi https://gephi.org/ (no usa python)

Grafo de un proyecto



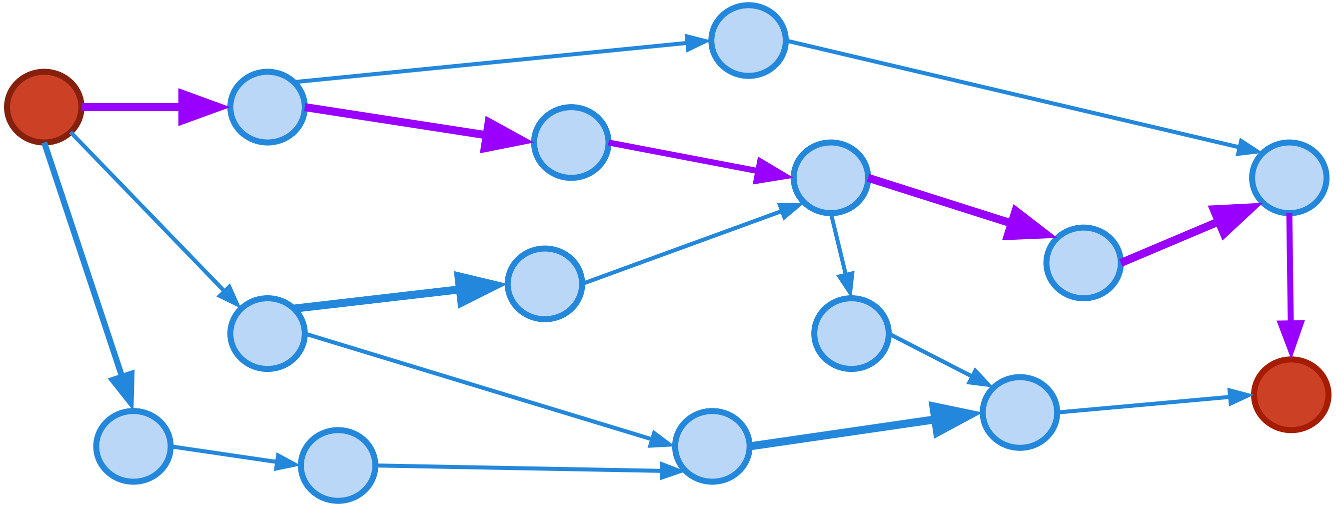
La red que utilizaremos para modelar los proyectos siempre tienen un nodo de inicio y un nodo de fin. Cada arco será una tarea. El **peso** (weight) de cada arco indicará su duración y en este ejemplo esta representada visualmente con su grosor. En la matriz de adyacencia los pesos de los arcos estan representados en los elementos de la matriz.

El camino crítico



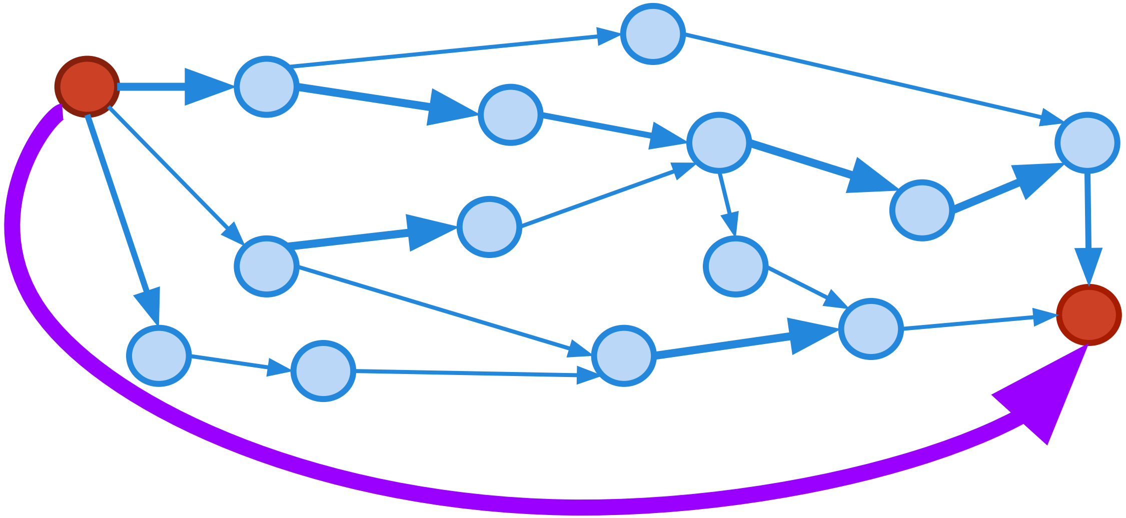
Con el algoritmo del Critical Path Method obtenemos el ***camino crítico*** (violeta).
Matematicamente: ¿qué representa el camino crítico?

El camino crítico: el camino mas largo



El camino crítico representa el camino mas largo entre dos nodos de interes en la red en términos de las unidades en la que estan pesados los arcos: duración de la tarea.

El camino crítico: el camino mas largo

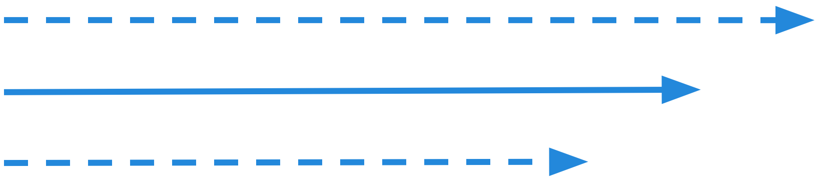


El camino mas largo no necesariamente tiene el mayor número de tareas. Si al grafo anterior le agregamos una nueva tarea que no tiene precedencias y cuya duración es mucho mayor que la duración del camino crítico del caso anterior, entonces esta nueva tarea pasa a ser el camino crítico y termina representando el camino mas largo entre los nodos de inicio y fin.

PERT

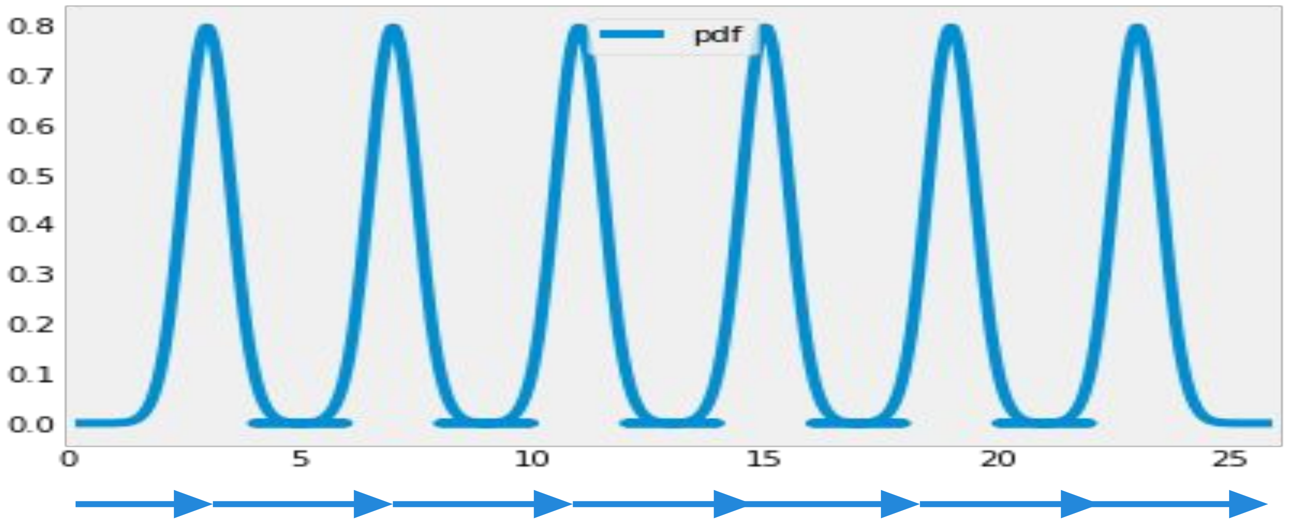
Program Evaluation and Review Technique

Grafos estocásticos



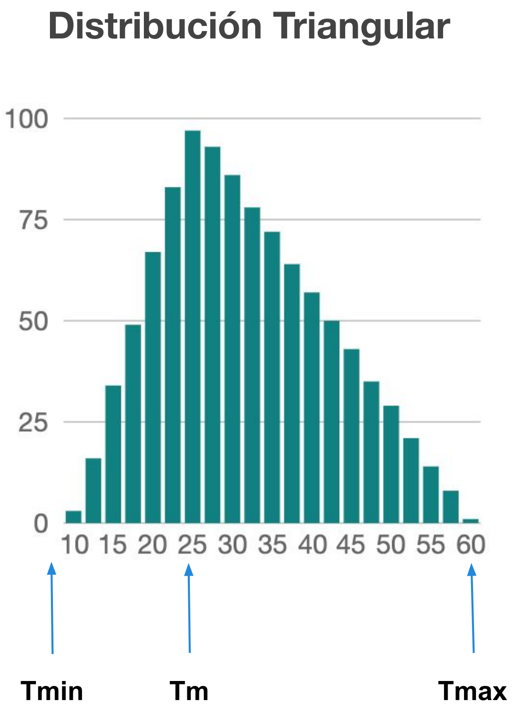
Existen distintos tipos de grafos estocásticos. Existen los casos donde la cantidad de arcos y nodos son aleatorios por ejemplo. De todos modos, en el caso de Planificación de Proyectos se supone que tanto la precedencia como la cantidad de tareas es una constante y son derministicos (no-aleatorios). Lo que si puede llegar a variar aleatoriamente bajo cierta distribución de probabilidad es la duración de las tareas. La pregunta natural sería ¿Cómo modelizar el camino crítico cuando las duraciones de las tareas se contemplan estocásticas? Es decir, cuando la duración de las tareas es una variable aleatoria.

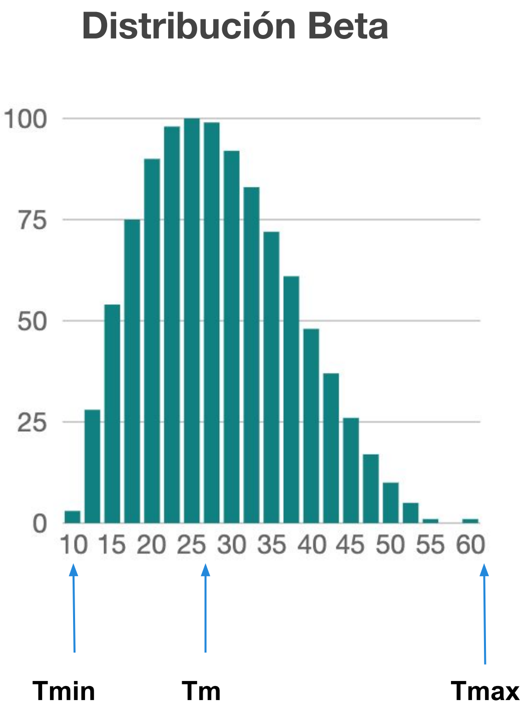
Grafos estocásticos



Si la duración de las tareas es estocástica entonces existirá una probabilidad de que el proyecto termine en fecha, otra probabilidad de que el proyecto termine antes de lo esperado y otra probabilidad de que el proyecto termine mas tarde de lo esperado.

Distribución de triangular y beta-PERT





Para modelizar la duración de las tareas se utilizará una distribución Beta. Esta distribución requiere de:

- Tiempo mínimo de duración de la tarea (**Tmin**).
- Tiempo máximo de duración de la tarea (**Tmax**).
- Tiempo medio/moda o mas probable de duración de la tarea (**Tm**).

Para simplificar cálculos se usará una distribución triangular que contiene los mismos parámetros que la beta.

Distribución Triangular

$$t_e = \frac{t_{\min} + 4t_m + t_{\max}}{6}$$

$$\sigma = \frac{t_{\max} - t_{\min}}{6}$$

Vamos a suponer que los tiempos mínimos, máximos y medios de ejecución de **una tarea** están disponibles. Con esos parámetros se podrá computar la media y el desvío standard de la distribución triangular.

Grafos estocásticos

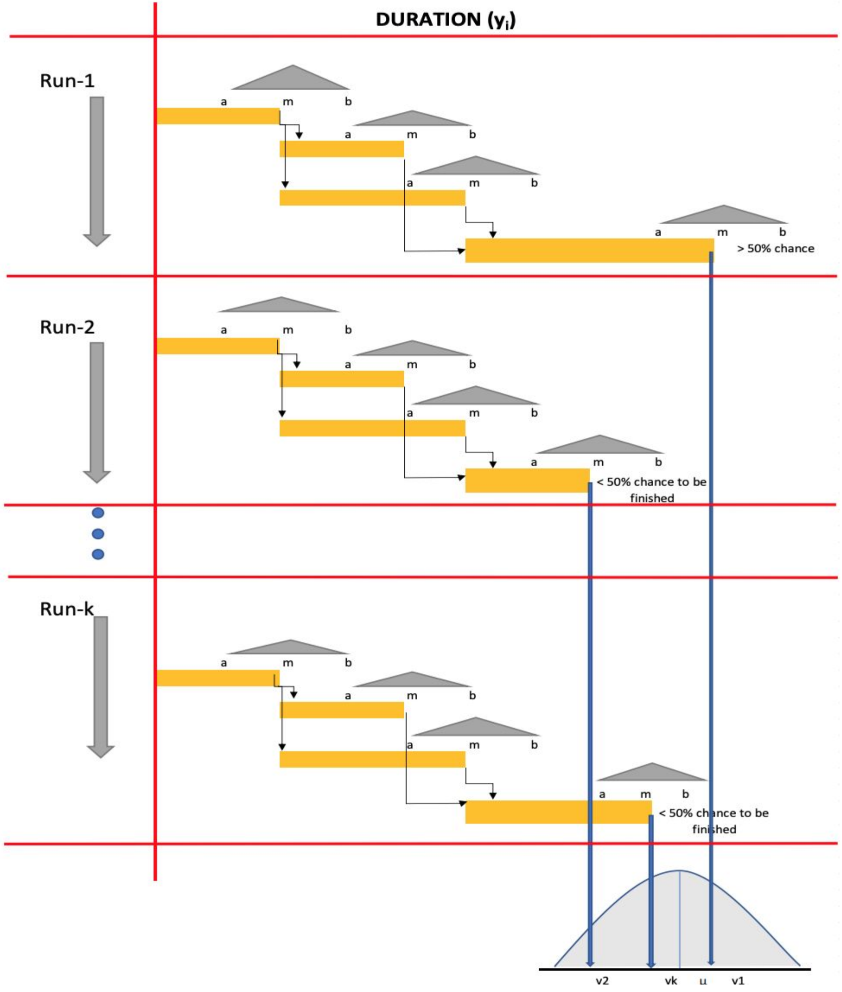
$$\mu = \sum t_e^{(cc)}$$

$$\sigma = \sqrt{\sum \sigma_{cc}^2}$$

La duración total **X** del proyecto será una variable aleatoria, dado que esta compuesto por la duración de las tareas que también lo son.

Haciendo uso del teorema central del límite, la suma de variables aleatorias aproxima a una distribución normal cuya media y varianza serán igual a la suma de las correspondientes a las tareas que la componen.

La media como el desvío standard del proyecto se calcularán únicamente a partir de los valores resultantes de las **tareas críticas**.



El camino crítico de la red se calculará utilizando los tiempos esperados ‘**te**’ de las tareas. Es decir, el peso de los arcos será ‘te’.

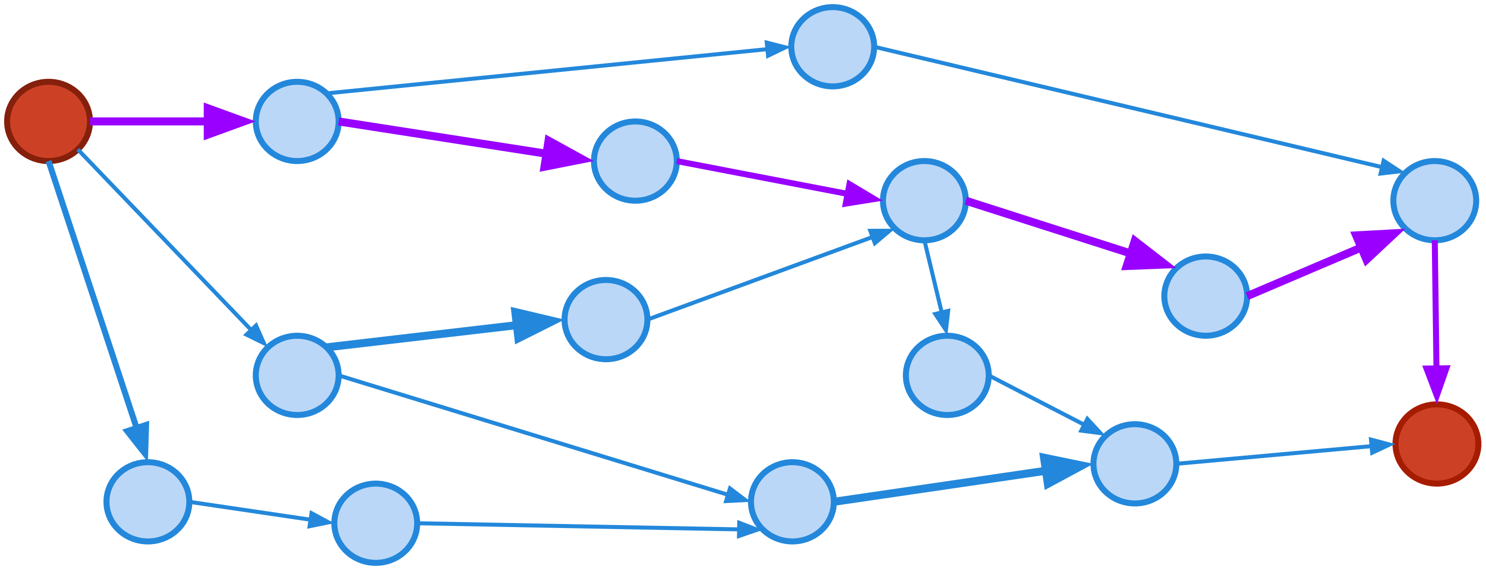
Una vez se que se tiene modelizado el camino crítico se obtiene una distribución de probabilidad sobre la duración del proyecto.

Con la distribución de probabilidad se puede simular la duracion del proyecto realizando un muestreo de monte-carlo/transformada inversa sobre dicha distribucion.

En cada experimento se obtendrá una duración distinta del proyecto.

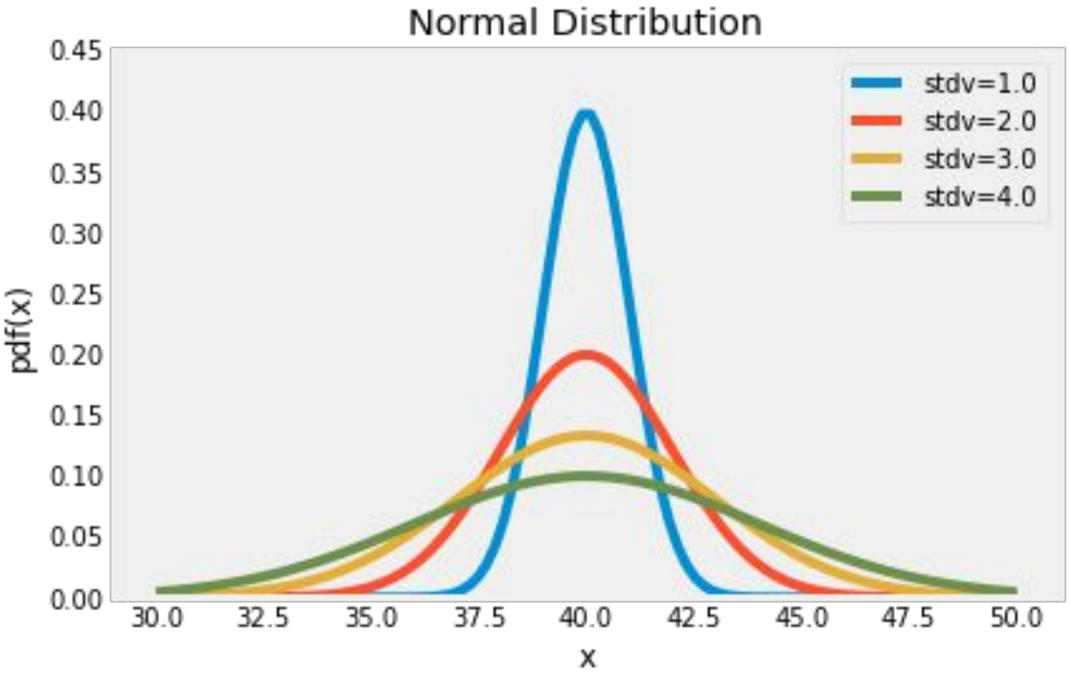
*Application of Monte Carlo simulation and PERT/CPM techniques in planning of construction projects: A Case Study

El camino crítico



Para calcular el camino crítico primero se debe obtener el tiempo esperado **te** de cada tarea. Luego el algoritmo para obtener que tareas pertenecen al camino crítico es equivalente al caso determinista.

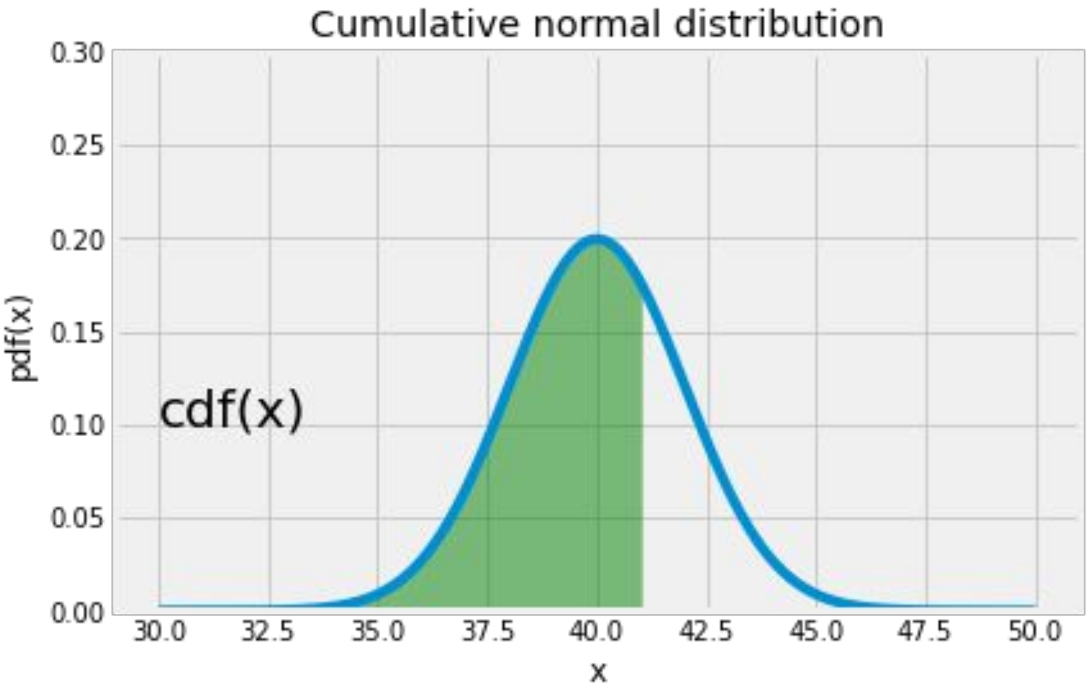
Varianza y Riesgo



Suponer que hay 4 versiones de un mismo proyecto evaluados por PERT. Los 4 proyectos tienen la misma duración esperada en 40 meses.

De todos modos hay diferencias significativas en la varianza. Cuanto mayor sea la varianza menos confianza vamos a tener para asegurar que el proyecto terminará a tiempo. Por esta razón los proyectos de menor varianza tendrán menos riesgo.

Probabilidad de finalización



Una vez que la duración del proyecto está caracterizada por una distribución normal gaussiana podemos preguntarnos cuál es la probabilidad de que el proyecto finalice antes de determinado tiempo. En la figura: Dado un determinado mu y sigma ¿cuál es la probabilidad de que el proyecto finalice antes de los 41 dias?

Probabilidad de finalización

$$x \sim N(\mu, \sigma)$$

$$Z = \frac{x - \mu}{\sigma}$$

$$P(x_i < X) = P(z_i < Z)$$

Suponemos X una variable aleatoria de distribución de densidad de probabilidad normal caracterizada por el mu y sigma del proyecto. Si transformamos a la distribución gaussiana a una normal-standard luego podemos preguntarnos la probabilidad acumulada para distintos valores de X (es decir, Z).