

# Connectivitat i Percolació\*

GRAU A      CURS 2020-2021

Departament de Ciències de la Computació  
alg@cs.upc.edu

## Resum

*Aquest projecte té com a objectiu un estudi experimental de la possible transició (o transicions) de fase del nombre de components connexes d'un graf sotmès a un procés de percolació.*

*El projecte es farà en grups de 3 persones (excepcionalment 2 o 4, sota autorització expressa). Els grups han de ser formats per estudiants matriculats al mateix subgrup de problemes. La **composició dels grups** s'haurà de comunicar a alg@cs.upc.edu abans del **14 d'octubre de 2020**.*

*Mantindrem un canal a l'espai Slack de l'assignatura per possibles consultes sobre l'enunciat i/o el desenvolupament del projecte. Consulteu-ho de tant en tant.*

*El lliurament de la pràctica es farà en línia via Racó, teniu temps fins les 10:00 hores del dia **26 d'octubre de 2020**.*

*Podria ser que alguns grups rebeu preguntes d'algun professor per e-mail o sigueu convocats a una entrevista cap a finals de novembre (una setmana abans publicariem la planificació de dates al Racó).*

## I. OBJECTIUS

L'objectiu principal d'aquesta pràctica és portar a terme un estudi experimental de la possible existència de transició (o transicions) de fase per una propietat de grafs quan les arestes i /o els nodes poden fallar. L'objectiu secundari és obtenir coneixements sobre diferents models de grafs aleatoris d'ús habitual en els estudis experimentals d'algorismes per a problemes sobre grafs amb fallides.

Les fallides permanents a un graf es poden modelitzar de diferents maneres. En aquest projecte podeu assumir que es corresponen a un senzill procés aleatori de percolació per nodes (*site percolation*) o per arestes (*bond percolation*). El model de percolació a un graf està definit per un paràmetre  $q \in [0, 1]$  que representa la probabilitat de fallida i proporciona un nou graf  $G_q$ . En un procés de percolació per nodes a un graf  $G$ , el paràmetre  $q$  representa la probabilitat que un node no falli.

---

\*La versió més actualitzada d'aquest document, així com qualsevol material addicional relacionat es publicarà al Racó.

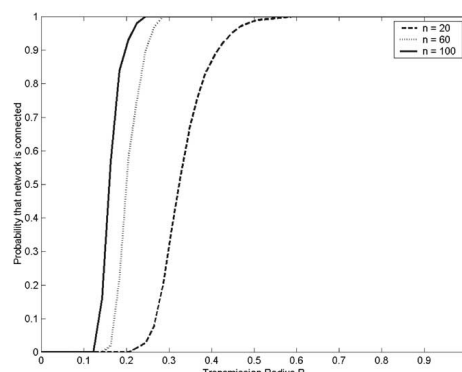
Així s'obté un graf  $G_q$  en què, per a cada node  $u \in V(G)$ , de forma independent, decidirem si el node  $u$  continua al graf (amb probabilitat  $p = 1 - q$  no falla) o si s'esborra  $u$  (amb probabilitat  $q$  falla). En un procés de percolació per arestes el paràmetre representa la probabilitat de fallida de una aresta. En aquest cas, per a cada aresta  $e \in E$ , mantindrem  $e$  amb probabilitat  $1 - q$  o la traurem amb probabilitat  $q$ . Observeu que  $V(G_q) = V(G)$  en el cas de percolació per arestes.

Una vegada tenim fixat un procés de percolació per a un graf concret  $G$ , volem estudiar per a quins valors de  $q$  podem esperar que certa propietat  $\Pi$  es doni al graf  $G_q$ . Per a moltes propietats  $\Pi$  es pot trobar un valor de  $q_\Pi$  de manera que, amb alta probabilitat, els grafs  $G_q$  amb  $q > q_\Pi$  verifiquen la propietat  $\Pi$ ; mentre que, amb alta probabilitat, els grafs  $G_q$  amb  $q < q_\Pi$  no verifiquen  $\Pi$ . Quan es dona aquest tipus de comportament es diu que la propietat  $\Pi$  presenta una *transició de fase* al voltant de  $q_\Pi$ .

Una descripció exacta d'un procés de *site percolation* a una graella quadrada  $n \times n$  el podeu trobar al material de suport del llibre *Algorithms* [1] (<https://algs4.cs.princeton.edu/lectures/keynote/15UnionFind-2x2.pdf>, a partir de la transparència 45). En concret, s'estudia la transició de fase de la propietat d'existència d'una connexió entre els nodes de la part de dalt amb els nodes de la part de sota d'una graella, sota percolació de nodes. En aquest cas se sap que la transició de fase per graelles grans apareix al voltant d'una probabilitat de percolació de 0.593.

En el cas de models parametritzats de grafs aleatoris la transició de fase estudia la variabilitat d'una propietat (o una mesura) en relació al valor (o valors) dels paràmetres que defineixen el graf. Com a exemple, considereu la propietat de la connectivitat de grafs (és a dir, que un graf sigui o no connex) aplicada a grafs aleatoris. Considereu que disposem d'una funció  $\text{Graf}(n, r)$  que permet generar un graf aleatori geomètric de  $n$  vèrtexs. En aquest model es seleccionen, amb distribució uniforme,  $n$  punts al quadrat unitari  $([0, 1] \times [0, 1])$  i es connecten dos nodes si la distància entre ells és menor o igual que  $r$ . Per a un valor de  $r$ , i fent un nombre raonable de crides, podem estimar la probabilitat que un graf amb  $n$  vèrtexs del model sigui connex.

La transició de fase, si n'hi ha, ens mostraria un valor del paràmetre  $r$  on la probabilitat de ser connex, per a valors  $r' < r$ , és radicalment diferent a la de valors  $r' > r$ . Com us podeu imaginar, la transició de fase no sempre es dona per a valors de  $n$  petits i per aquesta raó cal fer experiments que facin créixer el valor de  $n$  fins a tenir una certa evidència de l'existència de la transició de fase o la constatació que no apareix. A la figura de sota teniu una imatge que mostraria una transició de fase.



Us demanem que analitzeu experimentalment l'existència o no de transició (o transicions) de fase en relació al nombre de components connexes del graf percolat. Observeu que el nombre de components connexes no és un valor booleà i per tant poden aparèixer més d'un canvi de comportament al rang de valors de la probabilitat. Aquest estudi s'ha de realitzar sobre models parametritzats de grafs tant deterministes com aleatoris. Això vol dir que a més de les definicions de grafs parametritzats haureu d'estudiar els procediments que permeten generar un graf al model.

Tingueu en compte que volem analitzar el comportament asimptòtic en relació al nombre de nodes però que els models de grafs poden tenir altres paràmetres a més del nombre de nodes.

Per això us demanem que implementeu algorismes per a:

- (a) Donats  $G$  i  $q \in [0, 1]$  obtenir mostres dels grafs corresponents al procés de percolació de nodes i d'arestes amb probabilitat  $q$ .
- (b) Donat un graf  $G$  obtenir el nombre de components connexes que té.

Aquests algorismes els haureu de fer servir per a:

- (c) Estudiar la possible transició de fase a graelles quadrades  $n \times n$ , sota un procés de percolació de nodes i un d'arestes.
- (d) Estudiar la possible transició de fase en grafs geomètrics aleatoris connexos (Random geometric graphs), sota un procés de percolació d'arestes.
- (e) Estudiar la possible transició de fase sota un procés de percolació de nodes i un d'arestes a un altre model de grafs, per exemple graelles triangulars, graelles tridimensionals o un altre model de graf aleatori.

Tingueu en compte que per a poder veure experimentalment l'existència o no de transició (o transicions) de fase, haureu de treballar amb grafs grans, i potser caldrà que considereu casos en els quals la probabilitat del procés de percolació sigui també una funció dels paràmetres que defineixen el model de graf sota consideració.

Cal també que tingueu en compte la transició de fase de la connectivitat en grafs geomètrics aleatoris per tal de garantir que pels valors seleccionats el procés de percolació es porta a terme sempre sobre un graf connex.

D'altra banda, els grafs percolats són grafs aleatoris i per tant caldrà tenir resultats sobre una mostra prou gran per tal de garantir que els valors mitjans utilitzats a l'estudi són significatius.

## II. ENTREGA

El nivell de sofisticació i esforç dedicat a la pràctica és opcional i es tindrà en compte a l'hora d'avaluar-la. A la versió més senzilla (suficient per a aprovar el projecte si s'acompanya d'un bon disseny d'experiments) caldria que implementéssiu programes en C++ per als problemes proposats als apartats (a) i (b), i justificar els resultats experimentals per als models (c) i (d). Per optar a una bona nota caldria fer, a més, l'anàlisi de la transició de fase a un altre model (e).

## III. ALGUNS PUNTERS

Aquest document és intencionadament vague. Per tant, a més d'analitzar i experimentar amb diferents famílies de grafs, processos de percolació, grafs aleatoris i propietats, haureu de documentar les fonts d'informació, les decisions preses i el disseny d'experiments que els hi donen suport. Per a qualsevol font d'informació que feu servir cal que proporcioneu una referència adient, l'adreça web (si cal), les modificacions i/o simplificacions fetes en les vostres implementacions i/o la selecció de models de grafs.

Com a únic requisit us demanem que cobriu les implementacions per resoldre els problemes descrits als apartats (a) i (b). Per portar a terme l'estudi haureu de utilitzar generadors de grafs en els diferents models. Aquests generadors els podeu dissenyar vosaltres o fer servir implementacions ja fetes (en aquest cas cal que documenteu les fonts d'on heu tret els generadors).

Pel que fa als models de grafs aleatoris, podeu trobar alguns models per triar seguint els enllaços a [https://en.wikipedia.org/wiki/Random\\_graph](https://en.wikipedia.org/wiki/Random_graph) o les implementacions de generadors a NetworkX. També podeu trobar definicions d'alguns models de grafs aleatoris als capítols 12–15 del llibre [2].

## IV. QUÈ CAL LLIURAR

Cal lliurar una carpeta comprimida (.zip) que contingui:

- Una documentació adequada del algorismes i mètodes que heu implementat, les proves que heu fet i la comparació dels resultats que heu obtingut. També

és interessant que indiqueu altres idees que hagueu provat (encara que no hagin donat bons resultats) o d'altres que no heu explorat. El document en format PDF ha d'incloure les referències adients.

- Una carpeta amb totes les fonts necessàries per compilar i executar el vostre projecte (en C++). S'han d'incloure les instruccions per a la compilació i execució, així com els conjunts de dades utilitzats i/o els programes per generar-los.
- Tingueu en compte que la documentació entregada ens ha de permetre valorar el nivell d'assoliment de la competència transversal que hem d'avaluar: **Capacitat d'autoaprenentatge**. En el context del projecte hi han uns quants aspectes rellevants relacionats amb aquesta competència: propietats de grafs, transició de fase, models de grafs aleatoris, i el disseny i anàlisi dels experiments. La qualificació final del projecte reflectirà la qualitat del vostre aprenentatge, de l'experimentació feta i de la documentació on es reflectirà tot. La qualitat del codi entregat (programes) es pressuposa i representarà una part molt petita de la qualificació final.
- La documentació ha de recollir i presentar la feina feta, les fonts que s'han consultat, el que heu après i els resultats de l'experimentació. En particular és molt important que reflecteixi de forma succinta el que heu après. Si no es compleix aquesta condició, la qualificació final del projecte reflectirà la qualitat de la presentació i no es tindrà en consideració la resta de material lliurat.

## REFERÈNCIES

- [1] R. Sedgewick and K. Wayne. *Algorithms*. Addison-Wesley, 4th edition, 2011. <https://algs4.cs.princeton.edu/>
- [2] M.E.J. Newman. *Networks. An Introduction*. Oxford University Press, 2010.

