# Δρομολόγηση πακέτων σε δίκτυα υπολογιστών

Συμπληρωματικές σημειώσεις για το μάθημα
Αλγόριθμοι Επικοινωνιών

Ακαδημαϊκό έτος 2011-2012

## 1 Εισαγωγή

Οι παρακάτω σημειώσεις παρουσιάζουν την ανάλυση του άπληστου αλγορίθμου στο δίκτυο πεταλούδας και στο διδιάστατο πλέγμα για προβλήματα τυχαίων προορισμών καθώς και το σχεδιασμό και ανάλυση ενός πιθανοτικού αλγορίθμου δρομολόγησης μεταθέσεων για το δίκτυο πεταλούδας. Τα αποτελέσματα ισχύουν με μεγάλη πιθανότητα, δηλαδή, πιθανότητα τουλάχιστον $1 - 1/n^c$ όπου $n$ δείχνει το μέγεθος του δικτύου και $c$ είναι μια θετική σταθερά. Ξεκινάμε με την ανάλυση ενός γνωστού πειράματος.

## 2 Ένα ενδιαφέρον πείραμα

Έχουμε $n$ μπάλες και $n$ κουτιά. Για κάθε μπάλα, επιλέγουμε ομοιόμορφα τυχαία και ανεξάρτητα από τις υπόλοιπες ένα από τα $n$ κουτιά και την τοποθετούμε μέσα σε αυτό. Προφανώς, η διαδικασία είναι τέτοια ώστε μετά το τέλος του πειράματος κάθε κουτί περιέχει κατά μέσο όρο 1 μπάλα. Τι μπορεί να παρατηρήσουμε μετά από μια εκτέλεση του πειράματος; Για παράδειγμα, υπάρχει μια περίπτωσή όλες οι μπάλες να περιέχονται στο ίδιο κουτί. Επίσης, είναι 'πιθανό' κάθε κουτί να περιέχει ακριβώς μια μπάλα. Τι μπορούμε να ισχυριστούμε όμως σχετικά με το τι συμβαίνει συνήθως; Θα αποδείξουμε το παρακάτω λήμμα.

**Λήμμα 1** *Σε μια εκτέλεση του πειράματος, με μεγάλη πιθανότητα, κανένα κουτί δεν περιέχει περισσότερες από* $O\left(\frac{\log n}{\log\log n}\right)$ *μπάλες.*

**Απόδειξη:** Ας θεωρήσουμε ένα συγκεκριμένο κουτί. Για να έχει τουλάχιστον $r$ μπάλες, θα πρέπει να περιέχει κάποια $r$-άδα από τις $\left(\begin{array}{c} n \\ r \end{array}\right)$ δυνατές $r$-άδες $n$ μπαλών. Η πιθανότητα μια συγκεκριμένη $r$-άδα μπαλών να περιέχεται στο κουτί είναι $(1/n)^r$ (δηλ., η πιθανότητα κάθε μπάλα της $r$-άδας να περιέχεται στο κουτί). Οπότε, η πιθανότητα ότι το κουτί περιέχει κάποια από τις $\left(\begin{array}{c} n \\ r \end{array}\right)$ δυνατές $r$-άδες είναι το πολύ

$$\left(\begin{array}{c} n \\ r \end{array}\right)\left(\frac{1}{n}\right)^r.$$

Χρησιμοποιώντας την γνωστή ανισότητα $\left(\begin{array}{c} n \\ r \end{array}\right) \leq \left(\frac{ne}{r}\right)^r$, έχουμε ότι η πιθανότητα είναι το πολύ $\left(\frac{e}{r}\right)^r$.

Επιλέγοντας $r = 2\alpha e \frac{\log n}{\log\log n}$ για κάποια σταθερά $\alpha \geq 1$ έχουμε ότι η παράσταση γίνεται

$$\begin{aligned}
\left(\frac{e}{r}\right)^r &= 2^{r(\log e - \log r)} \\
&= 2^{2\alpha e \frac{\log n}{\log\log n}(\log e - 1 - \log\alpha - \log e - \log\log n + \log\log\log n)} \\
&\leq 2^{2\alpha e \frac{\log n}{\log\log n}(-1 - \log\log n + \log\log\log n)} \\
&\leq 2^{-2\alpha e \frac{\log n}{\log\log n}\frac{\log\log n}{2}} \\
&= 2^{-\alpha e \log n} \\
&= n^{-\alpha e}.
\end{aligned}$$

Σημειώνουμε ότι η πρώτη ανισότητα ισχύει επειδή $\alpha \geq 1$ ενώ η δεύτερη ανισότητα προκύπτει διότι $x \geq 2\log x - 2$ για κάθε $x > 0$.

Μέχρι στιγμής, έχουμε αποδείξει ότι η πιθανότητα ένα συγκεκριμένο κουτί να έχει τουλάχιστον $2\alpha e \frac{\log n}{\log\log n}$ μπάλες είναι το πολύ $1/n^{\alpha e}$. Οπότε, η πιθανότητα κάποιο κουτί να έχει τουλάχιστον $2\alpha e \frac{\log n}{\log\log n}$ μπάλες είναι το πολύ $1/n^{\alpha e - 1}$ και το λήμμα προκύπτει επιλέγοντας οποιαδήποτε σταθερά $\alpha \geq 1$. ■

# 3 Ανάλυση της συμπεριφοράς του άπληστου αλγορίθμου δρομολόγησης στο δίκτυο πεταλούδας

Μελετάμε τη συμπεριφορά του άπληστου αλγορίθμου στο δίκτυο πεταλούδας με $n$ κόμβους ανά επίπεδο (και διάστασης $\log n$) για τη δρομολόγηση προβλημάτων τυχαίων προορισμών. Σε ένα πρόβλημα δρομολόγησης τυχαίων προορισμών, κάθε κόμβος του πρώτου επιπέδου είναι αποστολέας ενός πακέτου. Ο προορισμός κάθε πακέτου επιλέγεται ομοιόμορφα τυχαία μεταξύ των $n$ κόμβων του τελευταίου επιπέδου ανεξάρτητα από τις επιλογές των άλλων πακέτων. Θα δείξουμε ότι ο αριθμός των πακέτων που θα περάσουν από οποιοδήποτε κόμβο είναι το πολύ $O\left(\frac{\log n}{\log\log n}\right)$, με μεγάλη πιθανότητα. Επομένως, με μεγάλη πιθανότητα, κανένα πακέτο δεν θα καθυστερήσει παραπάνω από $O\left(\frac{\log n}{\log\log n}\right)$ βήματα σε κάθε κόμβο της διαδρομής από τον αποστολέα στον προορισμό του και, επομένως, η δρομολόγηση θα ολοκληρωθεί σε το πολύ $O\left(\frac{\log^2 n}{\log\log n}\right)$ βήματα.

Σκεφτόμαστε ως εξής. Έστω ένας κόμβος $v$ του δικτύου πεταλούδας στο επίπεδο $i = 1, ..., \log n$. Τα πακέτα που είναι υποψήφια να περάσουν από τον κόμβο $v$ είναι αυτά που έχουν αποστολέα κάποιον κόμβο του επιπέδου $0$ από τον οποίο υπάρχει μονοπάτι που χρησιμοποιεί ο αλγόριθμος που να περνά από τον $v$. Από τον ορισμό του δικτύου πεταλούδας και τον ορισμό των μονοπατιών που χρησιμοποιεί ο άπληστος αλγόριθμος, γνωρίζουμε ότι υπάρχουν $2^i$ τέτοιοι κόμβοι και επομένως $2^i$ πακέτα που είναι υποψήφια να περάσουν από τον κόμβο $v$. Από αυτά τα υποψήφια πακέτα για να περάσουν από τον κόμβο $v$, θα περάσουν πραγματικά αυτά που θα επιλέξουν σαν προορισμούς κόμβους του τελευταίου επιπέδου τους οποίους μπορούμε να φτάσουμε από τον κόμβο $v$ ακολουθώντας μονοπάτια που χρησιμοποιεί ο άπληστος αλγόριθμος. Από τον ορισμό του δικτύου πεταλούδας και τον ορισμό των μονοπατιών που χρησιμοποιεί ο αλγόριθμος, γνωρίζουμε ότι υπάρχουν $2^{\log n - i}$ τέτοιοι κόμβοι. Αφού κάθε υποψήφιο πακέτο επιλέγει ισοπίθανα τον προορισμό του μεταξύ των $n$ κόμβων του τελευταίου επιπέδου, η πιθανότητα να περάσει από τον κόμβο $v$ είναι η πιθανότητα να διαλέξει κάποιον από αυτούς τους $2^{\log n - i}$ κόμβους του τελευταίου επιπέδου σαν προορισμό, δηλ., $2^{-i}$.

Έχουμε δηλαδή $2^i$ υποψήφια πακέτα να περάσουν από τον κόμβο $i$, καθένα από τα οποία έχει πιθανότητα $2^{-i}$ να περάσει όντως από τον κόμβο $v$. Όπως στην ανάλυση του πειράματος της προηγούμενης ενότητας, θα υπολογίσουμε ένα άνω φράγμα για την πιθανότητα να περάσουν τουλάχιστον $r$ υποψήφια πακέτα από τον κόμβο $v$. Παρατηρούμε ότι για να περάσουν τουλάχιστον $r$ υποψήφια πακέτα από τον κόμβο $v$, θα πρέπει να περάσει τουλάχιστον μια από τις $\left(\begin{array}{c} 2^i \\ r \end{array}\right)$ $r$-άδες από τα $2^i$ υποψήφια πακέτα από τον κόμβο $v$. Υπολογίζοντας αυτή την πιθανότητα και χρησιμοποιώντας την ίδια ανισότητα που χρησιμοποιήσαμε στην προηγούμενη παράγραφο, έχουμε ότι είναι το πολύ $\left(\frac{e}{r}\right)^r$. Χρησιμοποιώντας την ίδια ακριβώς λογική με αυτή που χρησιμοποιήσαμε στο τέλος της απόδειξης του Λήμματος 1, καταλήγουμε στο συμπέρασμα ότι η πιθανότητα να περάσουν τουλάχιστον $r = 2\alpha e \frac{\log n}{\log\log n}$ πακέτα από τον κόμβο $v$ είναι το πολύ $1/n^{\alpha e}$ (για $\alpha \geq 1$). Αφού έχουμε $n \log n$ κόμβους στα επίπεδα $1, ..., \log n$, η πιθανότητα να περάσουν τουλάχιστον $r = 2\alpha e \frac{\log n}{\log\log n}$ πακέτα από κάποιον από αυτούς είναι το πολύ $n \log n \cdot 1/n^{\alpha e} \leq 1/n^{\alpha e - 2}$ το οποίο είναι μικρότερο από $1/n^3$ αν επιλέξουμε $\alpha = 2$.

# 4 Ανάλυση της συμπεριφοράς του άπληστου αλγορίθμου δρομολόγησης στο διδιάστατο πλέγμα

Σε ένα πρόβλημα δρομολόγησης τυχαίων προορισμών στο διδιάστατο πλέγμα, κάθε κόμβος είναι αποστολέας ενός πακέτου. Ο προορισμός κάθε πακέτου επιλέγεται ομοιόμορφα τυχαία μεταξύ όλων των κόμβων του πλέγματος ανεξάρτητα από τις επιλογές των άλλων πακέτων. Για τον άπληστο αλγόριθμο στο διδιάστατο πλέγμα $n \times n$, δεν αναμένουμε να δείξουμε ότι δρομολογεί προβλήματα τυχαίων προορισμών σε χρόνο αρκετά μικρότερο από $2n - o(n)$ βήματα. Αυτό αποδεικνύεται στο παρακάτω λήμμα.

**Λήμμα 2** *Έστω ένα στιγμιότυπο δρομολόγησης τυχαίων προορισμών. Η πιθανότητα ότι ο αποστολέας κάθε πακέτου απέχει το πολύ* $2n - 2\sqrt{2n}\ln^{1/4} n$ *από τον προορισμό του είναι το πολύ* $1/n^2$.

**Απόδειξη:** Θεωρούμε δυο τρίγωνα στην πάνω αριστερή και κάτω δεξιά γωνία του πλέγματος πλευράς $\phi = \sqrt{2n}\ln^{1/4} n$. Για να απέχει ο αποστολέας κάθε πακέτου το πολύ $2n - 2\phi$ από τον προορισμό του, θα πρέπει να συμβαίνουν τουλάχιστον τα εξής: ο τυχαίος προορισμός κάθε πακέτου το οποίο ξεκινά από κόμβους στο πάνω αριστερά τρίγωνο να μην έχει προορισμό στο κάτω δεξιά τρίγωνο και αντιστρόφως. Η πιθανότητα να μη συμβεί αυτό για κάποιο συγκεκριμένο κόμβο μέσα στα τρίγωνα είναι $1 - \frac{\phi^2}{2n^2}$ εφόσον ο προορισμός κάθε πακέτου επιλέγεται ισοπίθανα μεταξύ των $n^2$ κόμβων και υπάρχουν $\phi^2/2$ κόμβοι σε κάθε τρίγωνο. Οπότε, η πιθανότητα κανένα πακέτο με αποστολέα κόμβο σε κάποιο τρίγωνο να μην έχει προορισμό κόμβο του τριγώνου που βρίσκεται απέναντι είναι

$$\begin{aligned}
\left(1 - \frac{\phi^2}{2n^2}\right)^{\phi^2} &= \left(1 - \frac{2\sqrt{\ln n}}{2n}\right)^{2n\sqrt{\ln n}} \\
&= \left(1 - \frac{\sqrt{\ln n}}{n}\right)^{\frac{n}{\sqrt{\ln n}}2\ln n} \\
&\leq e^{-2\ln n} \\
&= \frac{1}{n^2}.
\end{aligned}$$

∎

Μπορούμε να δείξουμε όμως ότι το μέγεθος της ουράς κάθε κόμβου είναι το πολύ $O\left(\frac{\log n}{\log\log n}\right)$, με μεγάλη πιθανότητα. Παρατηρούμε ότι το μέγεθος κάθε ουράς ενός κόμβου φράσσεται από πάνω από τον αριθμό των πακέτων που στρίβουν στον κόμβο. Οπότε, για να αποδείξουμε το επιθυμητό άνω φράγμα για το μέγεθος της ουράς, αρκεί να δείξουμε ότι η πιθανότητα ότι ο αριθμός των πακέτων που θα στρίψουν σε κάποιο κόμβο του πλέγματος είναι τουλάχιστον $r = 2\alpha e \frac{\log n}{\log\log n}$ είναι μικρή για κατάλληλη επιλογή της σταθεράς $\alpha$.

Η ανάλυση είναι παρόμοια με την ανάλυση των προηγούμενων παραγράφων. Θεωρούμε έναν κόμβο $v$ του πλέγματος. Υπάρχουν $n$ πακέτα στην ίδια γραμμή με τον κόμβο $v$ που είναι υποψήφια να στρίβουν σε αυτόν και από αυτά θα στρίψουν πράγματι μόνο όσα επιλέξουν σαν προορισμό κάποιον από τους κόμβους στην ίδια στήλη με τον $v$. Αφού οι προορισμού επιλέγονται ισοπίθανα μεταξύ των κόμβων, η πιθανότητα ένα υποψήφιο πακέτο να στρίβει στον κόμβο $v$ είναι $1/n$. Πάλι, μπορούμε να δούμε το πρόβλημα σαν μπάλες (τα υποψήφια πακέτα) που ρίχνονται ομοιόμορφα τυχαία σε κουτιά (οι στήλες) και, ακολουθώντας την ίδια ακριβώς λογική, μπορούμε να αποδείξουμε ότι η πιθανότητα τουλάχιστον $r$ πακέτα να στρίβουν στον κόμβο $v$ είναι το πολύ $1/n^{\alpha e}$. Οπότε, αφού έχουμε $n^2$ κόμβους, η πιθανότητα σε κάποιον από αυτούς να στρίβουν τουλάχιστον $r$ πακέτα είναι το πολύ $1/n^{\alpha e-2}$ το οποίο είναι μικρότερο του $1/n^3$ για $\alpha = 2$.

# 5 Ένας πιθανοτικός αλγόριθμος δρομολόγησης στο δίκτυο πεταλούδας

Μέχρι στιγμής, γνωρίζουμε ότι:

- Ο άπληστος αλγόριθμος δρομολόγησης στο δίκτυο πεταλούδας συμπεριφέρεται άσχημα στη χειρότερη περίπτωση για τη δρομολόγηση μεταθέσεων.
- Κανένας αλγόριθμος προκαθορισμένων μονοπατιών δεν μπορεί να πετύχει χρόνο δρομολόγησης μεταθέσεων καλύτερο από $\Omega(\sqrt{n})$ στη χειρότερη περίπτωση.
- Ο άπληστος αλγόριθμος συμπεριφέρεται πολύ καλά στη μέση περίπτωση (δηλαδή, σε προβλήματα τυχαίων προορισμών).

Μπορούμε να εκμεταλλευτούμε τη συμπεριφορά του άπληστου αλγορίθμου σε προβλήματα δρομολόγησης τυχαίων προορισμών και να σχεδιάσουμε τον παρακάτω γρήγορο αλγόριθμο δρομολόγησης μεταθέσεων:

Αρχικά κάθε κόμβος του επιπέδου $0$, υπολογίζει έναν ενδιάμεσο τυχαίο προορισμό του επιπέδου $\log n$. Η δρομολόγηση των πακέτων ακολουθεί τρεις φάσεις:

Φάση 1. Δρομολόγησε τα πακέτα στους ενδιάμεσους προορισμούς χρησιμοποιώντας τον άπληστο αλγόριθμο.

Φάση 2. Τα πακέτα γυρίζουν πίσω στο επίπεδο $0$ ακολουθώντας τις ευθείες συνδέσεις του δικτύου πεταλούδας.

Φάση 3. Δρομολόγησε τα πακέτα στους (πραγματικούς) προορισμούς χρησιμοποιώντας τον άπληστο αλγόριθμο.

Αναλύουμε την κάθε φάση ξεχωριστά.

**Φάση 1.** Έχουμε ένα πρόβλημα τυχαίων προορισμών. Ακολουθώντας την ανάλυση της ενότητας 3, έχουμε ότι η πιθανότητα να περάσουν από κάποιο κόμβο τουλάχιστον $\frac{2\alpha e \log n}{\log \log n}$ πακέτα είναι το πολύ $1/n^{\alpha e-2}$.

**Φάση 2.** Από τη στιγμή που ένα πακέτο ξεκινήσει να κινείται, δεν καθυστερείται από άλλα πακέτα μέχρι να φτάσει στο επίπεδο $0$. Ένα πακέτο μπορεί να καθυστερήσει το πολύ $\frac{2\alpha e \log n}{\log \log n} - 1$ βήματα στο επίπεδο $\log n$ (αφού έχουμε υποθέσει ότι το πολύ τόσα πακέτα μπορεί να περάσουν από κάποιον κόμβο του τελευταίου επιπέδου στη φάση 1) και στη συνέχεια ακολουθεί $\log n$ συνδέσεις μέχρι να φτάσει στον αντίστοιχο κόμβο του επιπέδου $0$.

**Φάση 3.** Η ανάλυση είναι αντίστοιχη με την ανάλυση της Φάσης 1 εξετάζοντας αντίθετα τα μονοπάτια των πακέτων από τους προορισμούς προς τους κόμβους του επιπέδου $0$ από όπου ξεκινούν στην αρχή της Φάσης 3. Θεωρούμε έναν κόμβο $v$ του επιπέδου $i = 1, ..., \log n$. Τα πακέτα που είναι υποψήφια να περάσουν από τον κόμβο $v$ είναι αυτά που έχουν σαν τελικό προορισμό κάποιον από τους $2^{\log n - i}$ κόμβους του επιπέδου $\log n$ τους οποίους μπορούμε να φτάσουμε από τον $v$ ακολουθώντας μονοπάτια που χρησιμοποιεί ο άπληστος αλγόριθμος. Από αυτά τα πακέτα, στη φάση 3, θα περάσουν πράγματι από τον κόμβο $v$ όσα έχουν βρεθεί στην αρχή της φάσης 3 σε κόμβους του επιπέδου $0$ από τους οποίους μπορούμε να φτάσουμε στον κόμβο $v$ ακολουθώντας μονοπάτια του άπληστου αλγορίθμου. Η ανάλυση συνεχίζεται όπως στις προηγούμενες ενότητες. Υποθέτοντας ότι στην αρχή της Φάσης 3, κανένας κόμβος του επιπέδου $0$ δεν περιέχει περισσότερα από $\frac{2\alpha e \log n}{\log \log n}$ πακέτα, μπορούμε να δείξουμε ότι η πιθανότητα να περάσουν από κάποιο κόμβο τουλάχιστον $\frac{2\alpha e \log n}{\log \log n}$ πακέτα είναι το πολύ $1/n^{\alpha e-2}$.

Οπότε, με πιθανότητα τουλάχιστον $1 - 2/n^{\alpha e-2} \geq 1 - 1/n^{\alpha e-3}$, κανένα πακέτο δεν καθυστερεί σε κάποιο κόμβο περισσότερα από $\frac{2\alpha e \log n}{\log \log n}$ βήματα κατά τις Φάσεις 1 και 3, ενώ η Φάση 2 χρειάζεται το πολύ $\log n + \frac{2\alpha e \log n}{\log \log n}$ βήματα. Η συνολική διαδρομή που έχουν να διανύσουν τα πακέτα στις Φάσεις 1 και 3 είναι $2 \log n$. Επιλέγοντας $\alpha = 2$, καταλήγουμε στο συμπέρασμα ότι ο αλγόριθμος δρομολογεί οποιαδήποτε μετάθεση σε χρόνο το πολύ $O\left(\frac{\log^2 n}{\log \log n}\right)$, με πιθανότητα τουλάχιστον $1 - 1/n^2$.