Αλγοριθμικές Μέθοδοι Βελτιστοποίησης με Έμφαση σε Κατανεμημένα Προβλήματα

Βασικές Τεχνικές Κατανεμημένης Βελτιστοποίησης (μέρος 3)

Δημήτρης Αμπελιώτης

Επίκουρος Καθηγητής, Ιόνιο Πανεπιστήμιο

Περιεχόμενα

- Το πρόβλημα της συμφωνίας ως προς το μέσο όρο (average consensus)
- Ο Αλγόριθμος Random Pairwise Gossip
- Ο Αλγόριθμος Path Averaging
- Γενικοί κανόνες ανανέωσης
- Ο κατανεμημένος υπολογισμός μέσου όρου ως ένα πρόβλημα βελτιστοποίησης
- Κατευθυνόμενα γραφήματα και ο αλγόριθμος Uniform Gossip

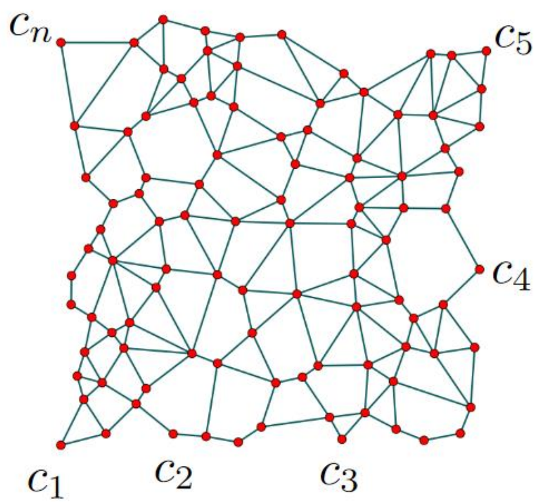
Το Πρόβλημα της Συμφωνίας ως προς το Μέσο Όρο

(Average Consensus)

Average Consensus

- Σε ένα **μη-κατευθυνόμενο** και **συνεκτικό** γράφημα, έχουμε
  - n κόμβους / agents
  - m ακμές (που αναπαριστούν δυνατότητα άμεσης επικοινωνίας)
  - Κάθε κόμβος έχει μια **τοπική** τιμή c_i
- Σκοπός είναι η εύρεση ενός **κατανεμημένου αλγορίθμου** για τον υπολογισμό του μέσου όρου

$$\bar{c} = \frac{c_1 + c_2 + \cdots + c_n}{n}$$



Average Consensus

- Γιατί να ασχοληθούμε με το πρόβλημα αυτό;
  - Η μελέτη ενός απλού προβλήματος κατανεμημένου υπολογισμού ελπίζουμε να μας βοηθήσει να κατανοήσουμε τους **θεμελιώδεις περιορισμούς** που προκύπτουν και να γνωρίζουμε τι πρέπει να προσέξουμε κατά τη μελέτη και εφαρμογή σε πιο σύνθετα προβλήματα
  - Αν και απλό, το πρόβλημα του κατανεμημένου υπολογισμού του μέσου όρου **εμφανίζεται ως υπό-πρόβλημα** σε αρκετά πιο σύνθετα προβλήματα όπως
    - Σε προβλήματα κατανεμημένης εκτίμησης (distributed estimation)
    - Σε προβλήματα κατανεμημένης ιχνηλάτησης (distributed tracking) όπου χρησιμοποιείται το κατανεμημένο φίλτρο Kalman
    - Σε προβλήματα κατανεμημένης μάθησης (**Federated Learning**)
    - Σε προβλήματα κατανεμημένης βαθμονόμησης (distributed sensor calibration), όπως π.χ. σε προβλήματα συγχρονισμού των ρολογιών των κόμβων ενός δικτύου
    - Σε προβλήματα κατανεμημένου σχηματισμού σμήνους (distributed formation control)
    - Σε μοντέλα συμπεριφοράς ομάδων ζώων / σε κοινωνικά δίκτυα (animal flocking and herding / opinion dynamics)

Average Consensus

- Διάφορες παραλλαγές του προβλήματος
  - Απλό σενάριο: Έχουμε ένα σταθερό (χρονικά αμετάβλητο) γράφημα, όπου οι γειτονικοί κόμβοι επιτρέπεται να ανταλλάσσουν μηνύματα με σύγχρονο τρόπο.
  - Ατελής επικοινωνία: Το περιεχόμενο των μηνυμάτων που ανταλλάσσονται αλλοιώνεται (εισαγωγή θορύβου, κβάντιση τιμών, ύπαρξη καθυστερήσεων)
  - Χρονικά μεταβαλλόμενη τοπολογία: Το γράφημα είναι χρονικά μεταβαλλόμενο, για παράδειγμα κάθε ακμή εμφανίζεται με μια πιθανότητα. Μοντελοποιούνται προβλήματα στην επικοινωνία. Σε ορισμένες περιπτώσεις χρησιμοποιούνται πιο «χαλαροί» ορισμοί για τη συνεκτικότητα των τυχαίων αυτών γραφημάτων.
  - Χρονικά μεταβαλλόμενη τοπολογία στη βάση γεωμετρικών χαρακτηριστικών: Το γράφημα μπορεί να είναι χρονικά μεταβαλλόμενο επειδή έχουμε κινούμενους κόμβους (drones, mobile robots) και η ύπαρξη μιας ακμής εξαρτάται από την απόσταση ανάμεσα σε αυτό το ζεύγος κόμβων
  - Σύγκλιση σε πεπερασμένο χρόνο: Αν και το βασικό κομμάτι της έρευνας ασχολείται με αλγορίθμους που συγκλίνουν (θεωρητικά) μετά από άπειρες επαναλήψεις, τα τελευταία χρόνια έχει υπάρξει έντονο ενδιαφέρον για αλγορίθμους που τερματίζουν σε πεπερασμένο αριθμό βημάτων

Average Consensus

- Βιβλιογραφία:
  - Classic book: Bertsekas and Tsitsiklis, Parallel and distributed computation: Numerical methods, Prentice Hall, 1989
  - Classic book (computer science point of view): Lynch, Distributed algorithms, Morgan Kaufmann, 1997
  - Seminal paper (1): Olfati-Saber, Murray, Consensus problems in networks of agents with switching topology and time delays, IEEE TAC, 2004
  - Seminal paper (2): Moreau, Stability of multi-agent systems with time-dependent communication links, IEEE TAC, 2005
  - Book on mobile agents coordination: Bullo, Cortés, Martínez, Distributed Control of Robotic Networks, Princeton, 2009
  - Survey on consensus in distributed estimation or control: Garin, Schenato, A survey on distributed estimation and control applications using linear consensus algorithms, in Networked Control Systems, Springer LNCIS, 2011
  - Survey on gossip: Dimakis, Kar, Moura, Rabbat, Scaglione, Gossip algorithms for distributed signal processing, Proc. of the IEEE, 2011
  - Survey on opinion dynamics: Acemoglu, Ozdaglar, Opinion dynamics and learning in social networks, Dynamic Games and Applications, 2011

Ο Αλγόριθμος Random Pairwise Gossip

Ο Αλγόριθμος Random Pairwise Gossip

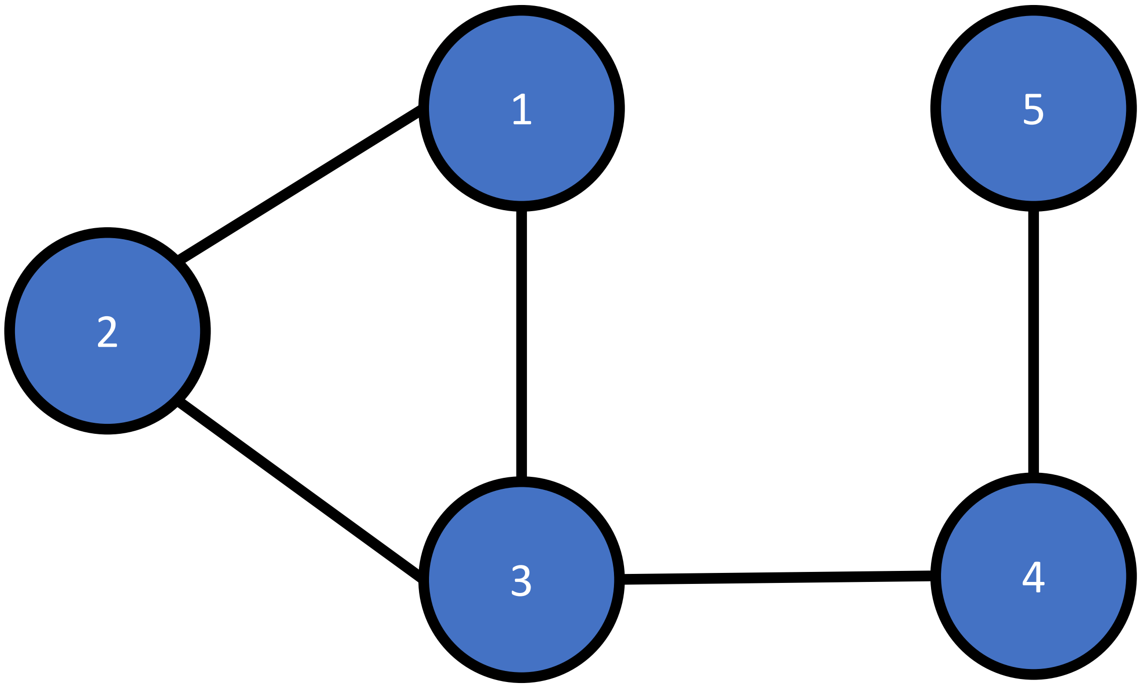
- Stephen Boyd, Arpita Ghosh, Balaji Prabhakar and Devavrat Shah, Randomized Gossip Algorithms, *IEEE Transactions on Information Theory,* 52.6 (2006), pp. 2508–2530

Random Pairwise Gossip

- Κάθε κόμβος i διατηρεί την εκτίμηση x_i^t για το μέσο όρο, κατά την επανάληψη / φάση t του αλγορίθμου
- Σε κάθε επανάληψη / φάση του αλγορίθμου, επιλέγουμε τυχαία μια ακμή
- Οι κόμβοι που συνδέονται από αυτή την ακμή επικοινωνούν και υπολογίζουν το μέσο όρο των τοπικών εκτιμήσεών τους
- Οι κόμβοι αυτοί ανανεώνουν τις τοπικές εκτιμήσεις τους θέτοντάς τες ίσες με το μέσο όρο που υπολόγισαν
- Οι υπόλοιποι κόμβοι απλά διατηρούν τις παλιές εκτιμήσεις τους

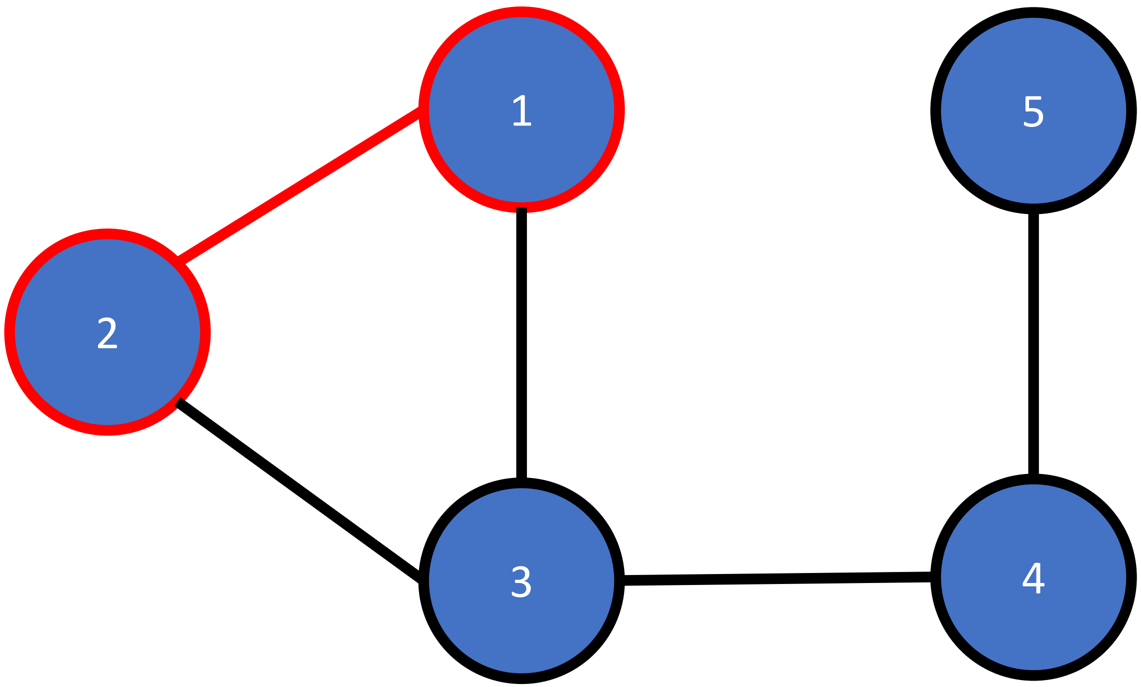
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



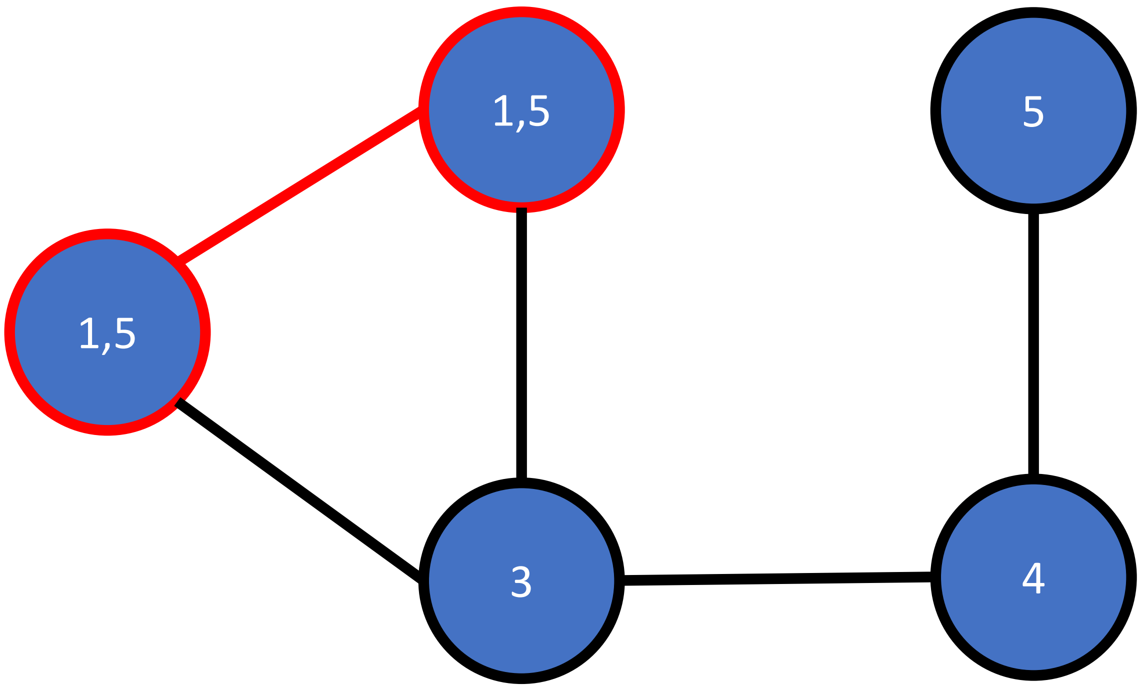
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



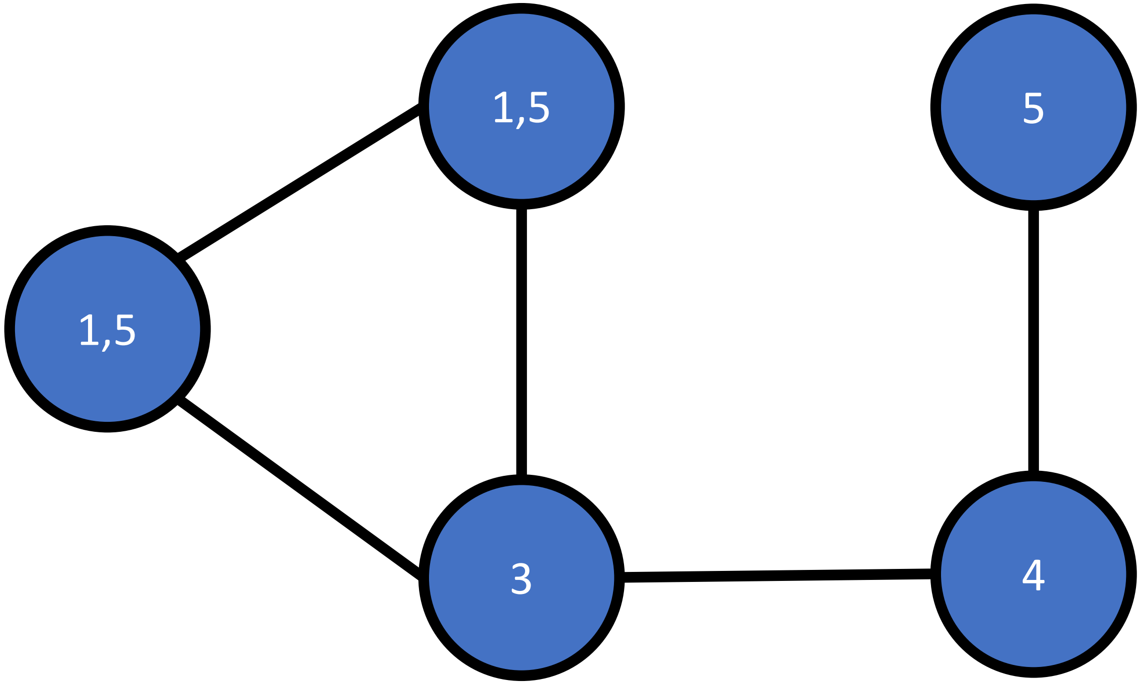
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



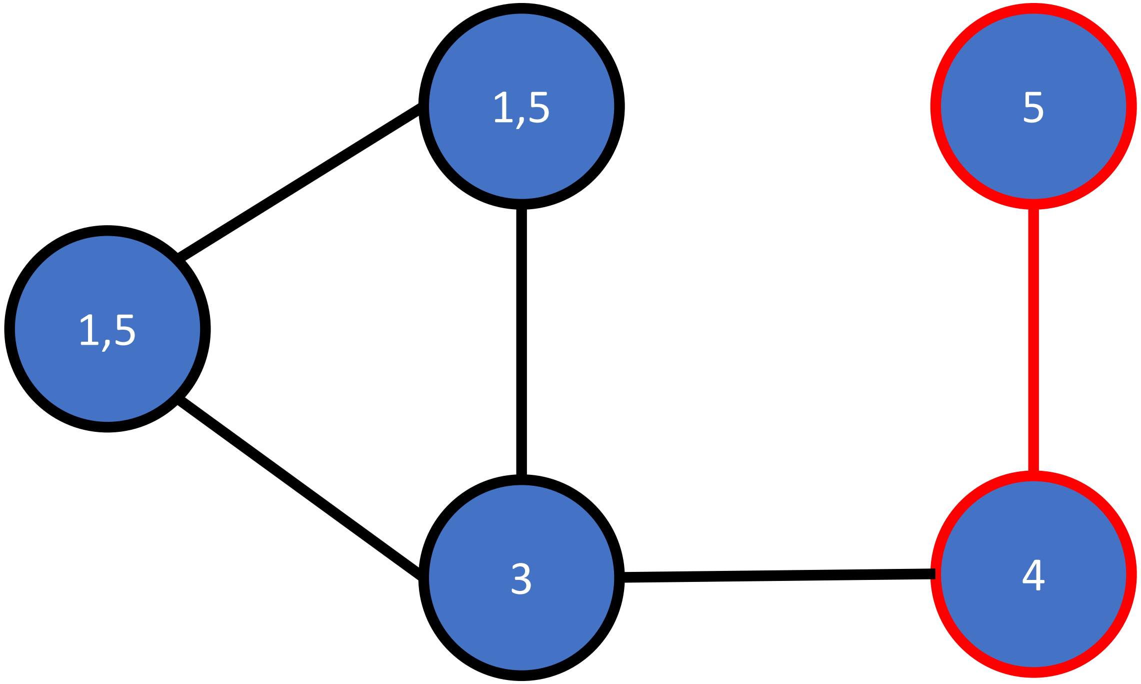
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



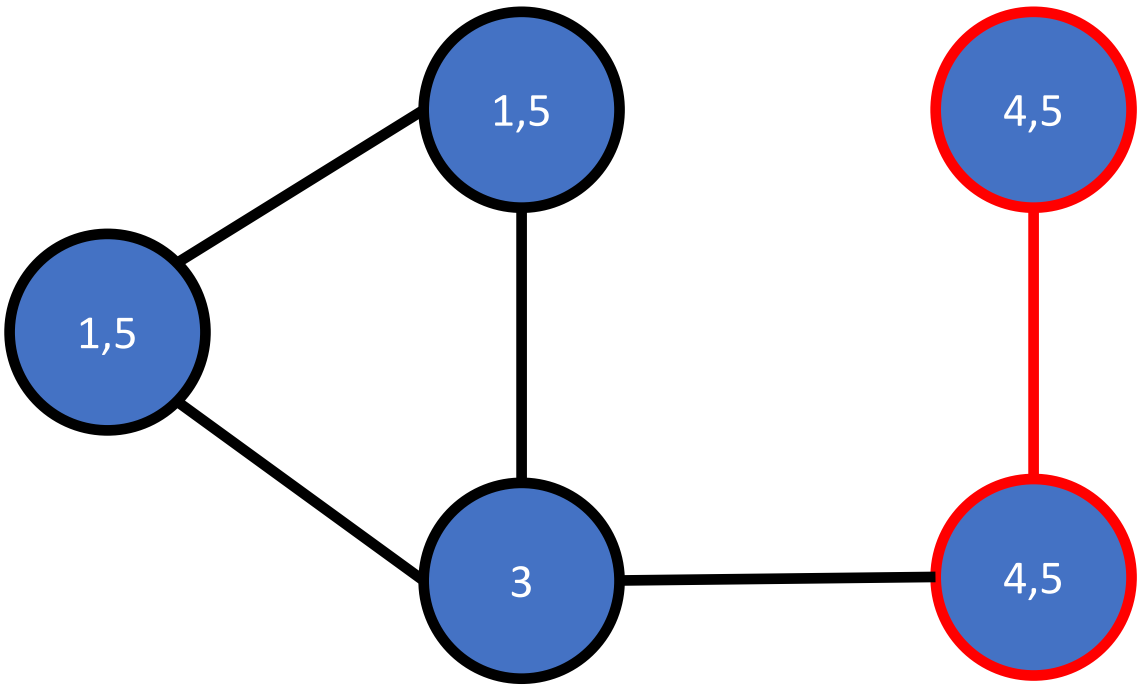
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



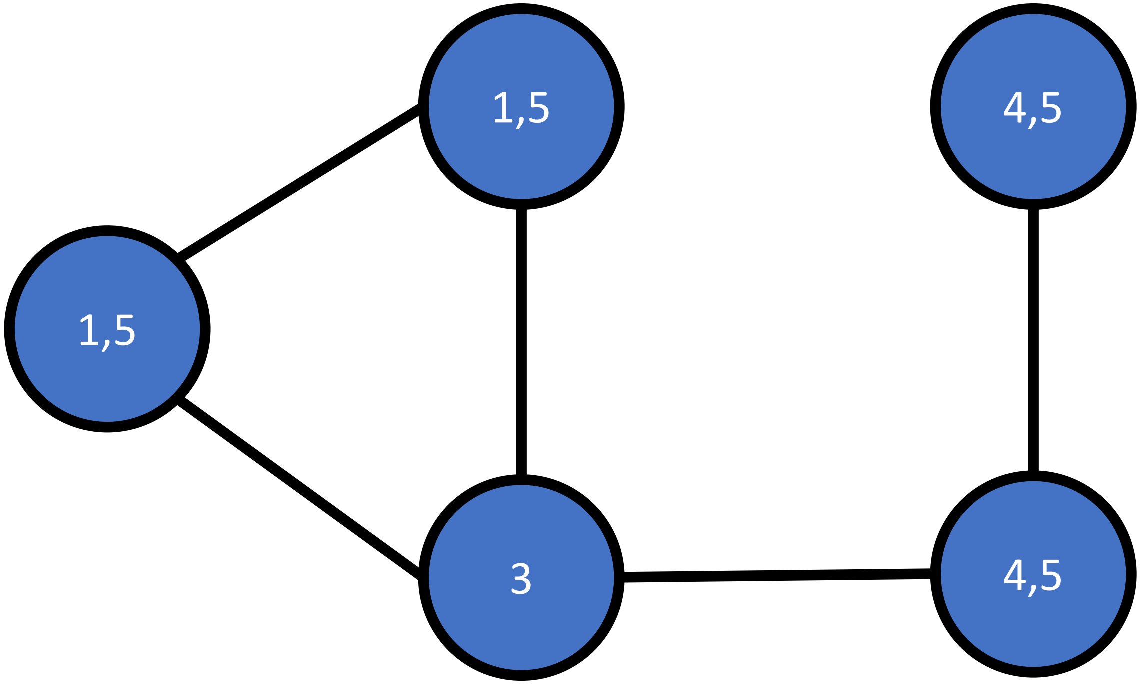
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



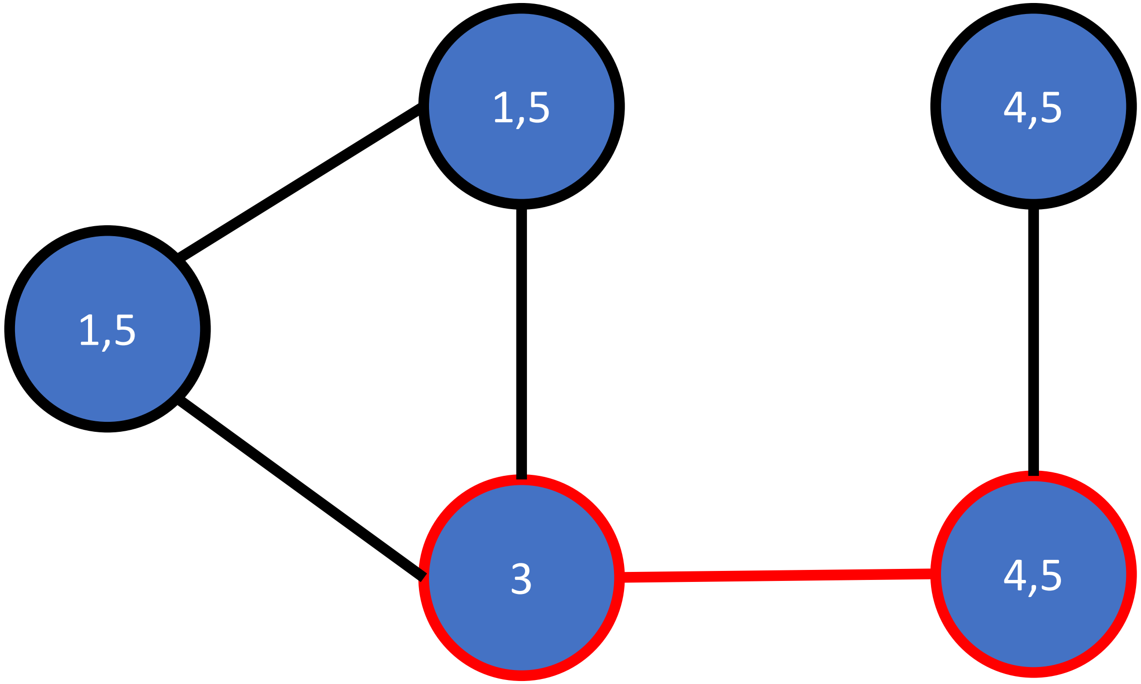
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



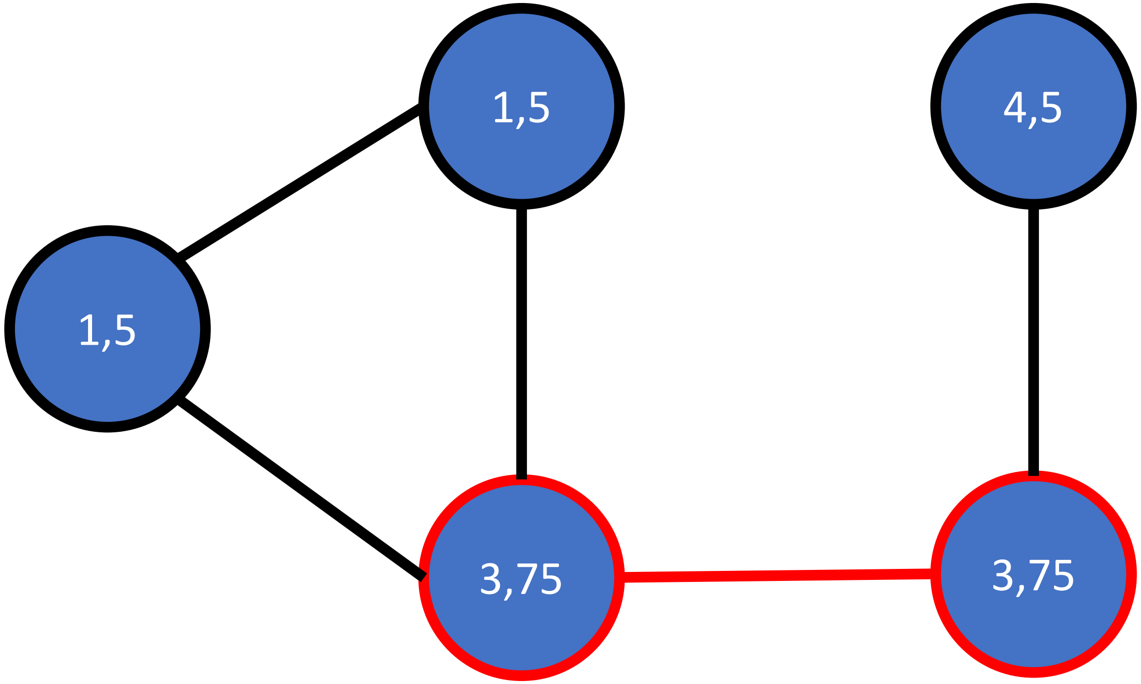
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



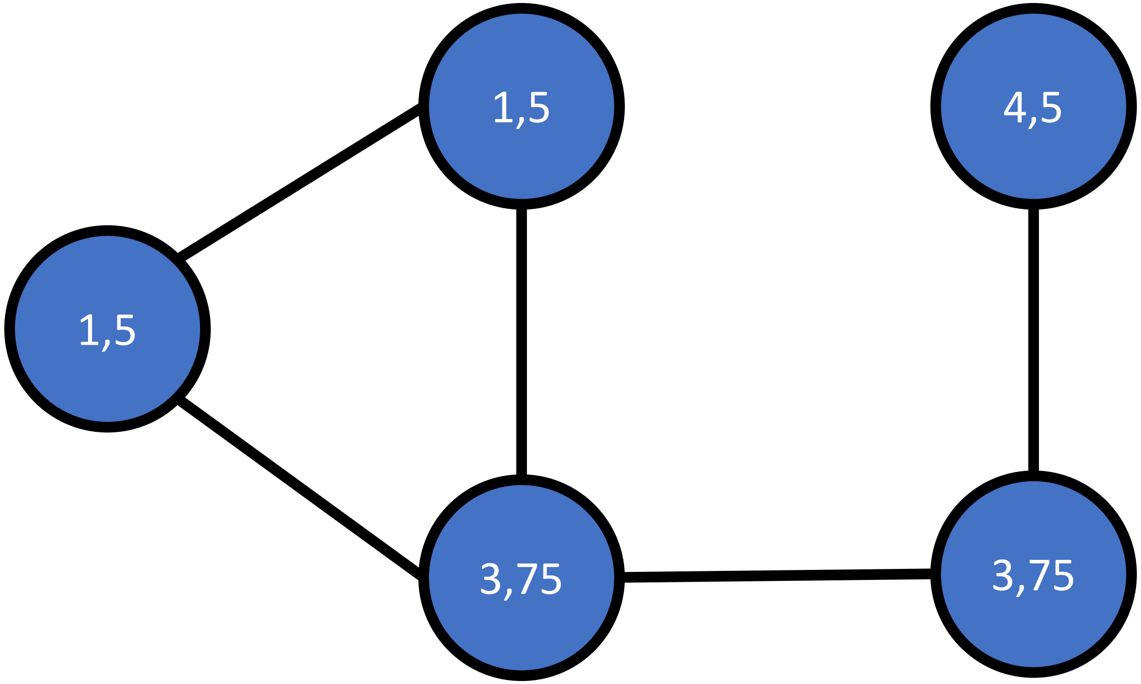
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



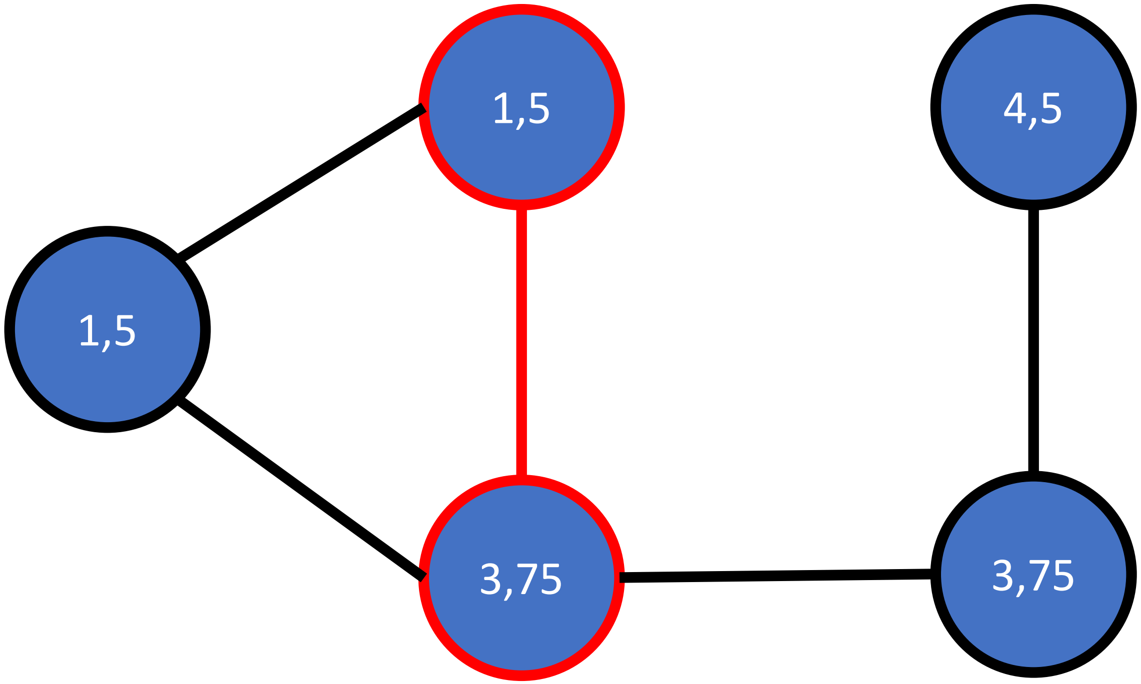
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



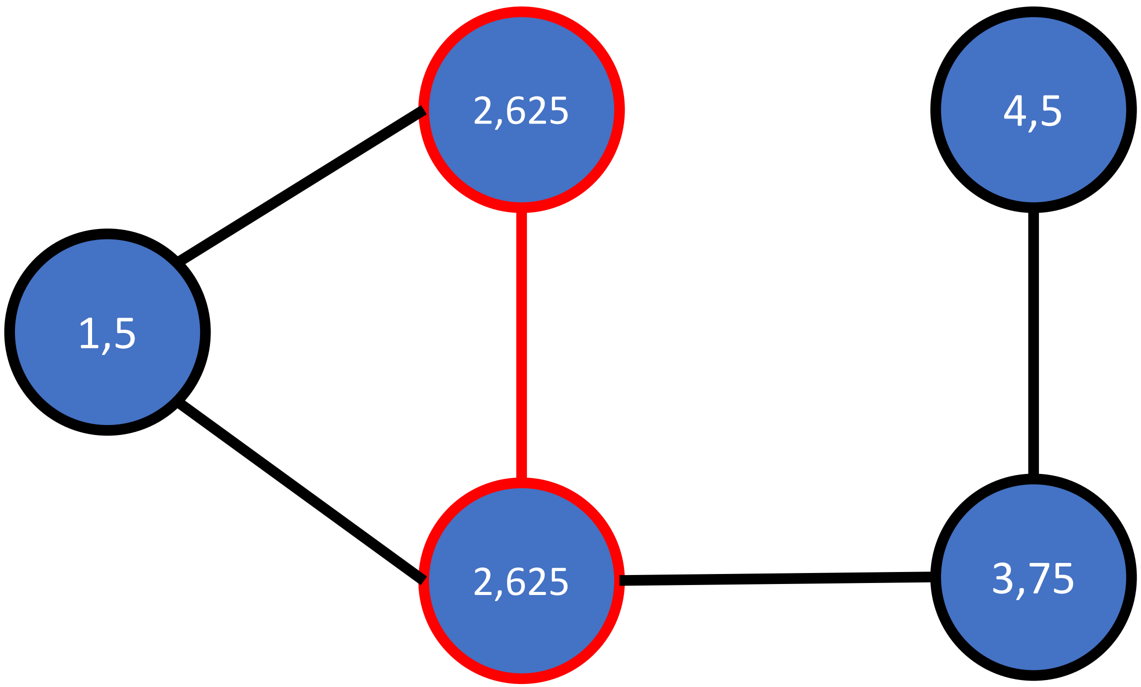
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



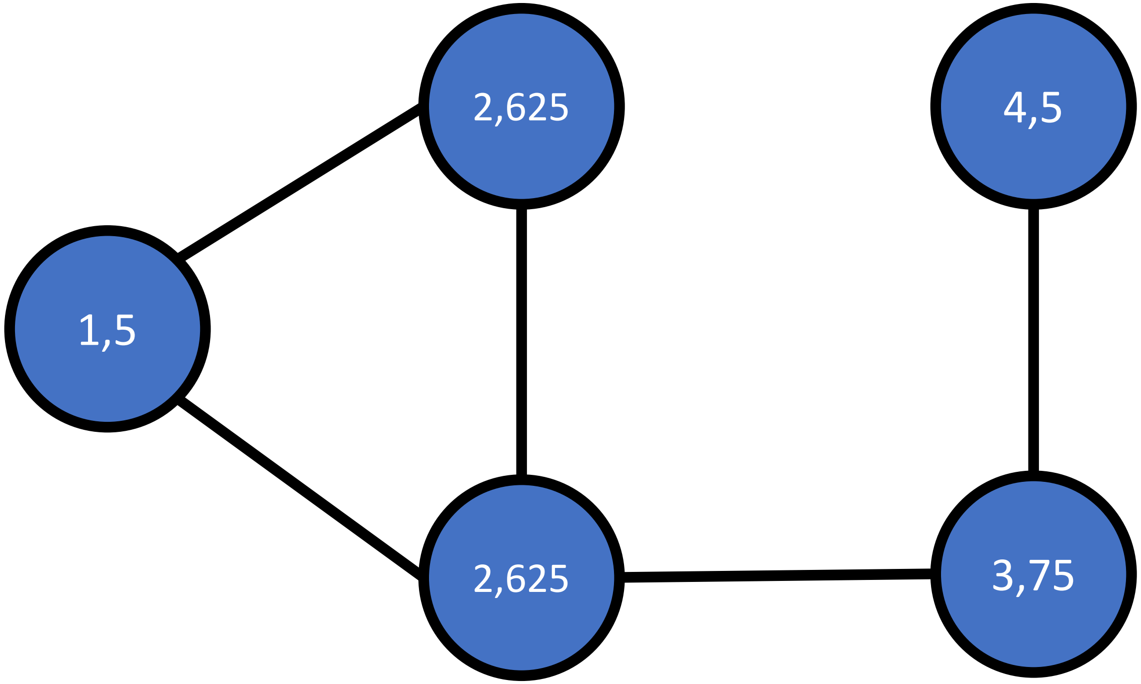
Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



Ο Αλγόριθμος Random Pairwise Gossip

- Παράδειγμα:



Ο Αλγόριθμος Random Pairwise Gossip

1. Initialize $x_i^0 = c_i$ for all $i = 1, 2, \ldots, n$
2. For $t \geq 0$ iterate:
   (a) Pick a random edge (i, j)
   (b) Set $x_i^{t+1} \leftarrow \frac{x_i^t + x_j^t}{2}$
   (c) Set $x_j^{t+1} \leftarrow \frac{x_i^t + x_j^t}{2}$
   (d) Set $x_u^{t+1} = x_u^t$ for all $u \notin \{i, j\}$

Ο Αλγόριθμος Random Pairwise Gossip

- Ο Αλγόριθμος Random Pairwise Gossip
  - Απλός
  - Εκτελεί τοπικές ανανεώσεις
  - Δεν απαιτεί γνώση για όλο το δίκτυο
  - Δεν απαιτεί κάποια προηγούμενη οργάνωση των κόμβων

- Ωστόσο, έχει αργή σύγκλιση
  - F. Bénézit, A. G. Dimakis, P. Thiran and M. Vetterli, "Order-Optimal Consensus Through Randomized Path Averaging," in *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 5150-5167, Oct. 2010
  - Στην εργασία αυτή έδειξαν πως για ένα μοντέλο τυχαίων γραφημάτων (random geometric graphs) ο αλγόριθμος Pairwise Gossip απαιτεί $\Theta(n^2)$ μηνύματα, ενώ προτείνουν ένα νέο αλγόριθμο ο οποίος απαιτεί μόλις $\Theta(n\log(n))$ μηνύματα

Ο Αλγόριθμος Path Averaging

Ο Αλγόριθμος Path Averaging

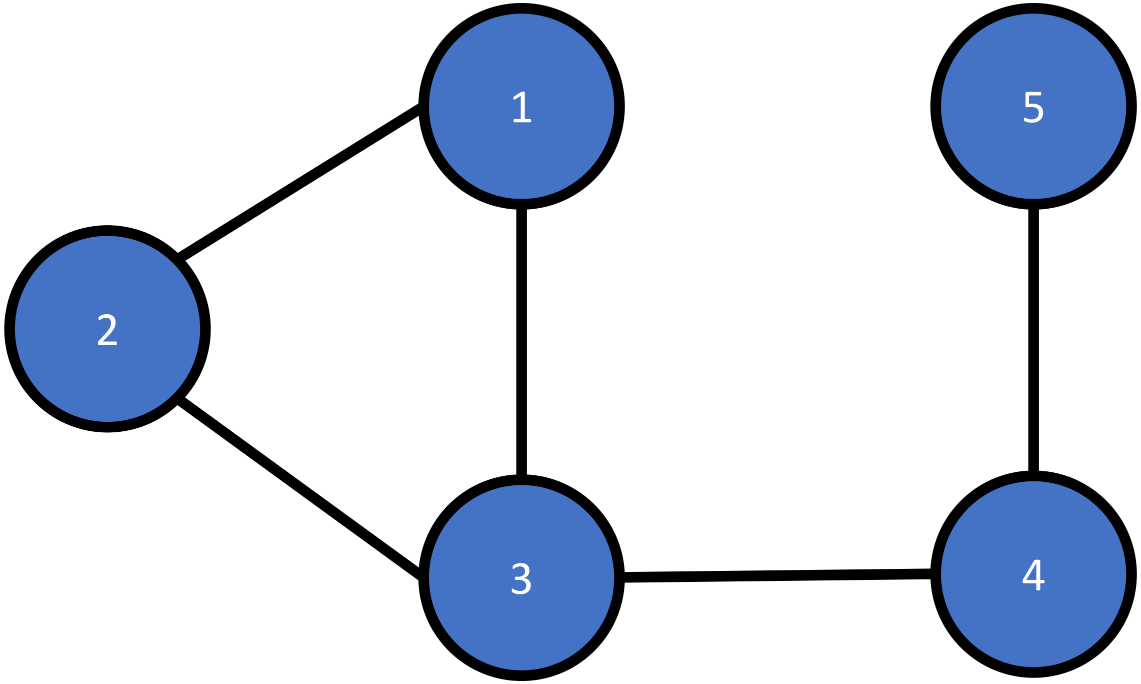
- Κερδίζει σε ταχύτητα σύγκλισης
- Χάνει επειδή απαιτεί «περισσότερη οργάνωση» των κόμβων

Path Averaging

- Κάθε κόμβος i διατηρεί την εκτίμηση x_i^t για το μέσο όρο, κατά την επανάληψη / φάση t του αλγορίθμου
- Σε κάθε επανάληψη / φάση του αλγορίθμου, επιλέγεται με τυχαίο τρόπο ένας κόμβος και δημιουργεί ένα τυχαίο μονοπάτι ξεκινώντας από αυτόν
- Οι κόμβοι κατά μήκος του τυχαίου μονοπατιού υπολογίζουν το άθροισμα των τιμών τους
- Ο τελευταίος κόμβος στο μονοπάτι υπολογίζει το μέσο όρο
- Το μήνυμα με το μέσο όρο μεταδίδεται προς τα πίσω στο μονοπάτι
- Οι κόμβοι του μονοπατιού ανανεώνουν τις εκτιμήσεις τους θέτοντάς τες ίσες με το μέσο όρο

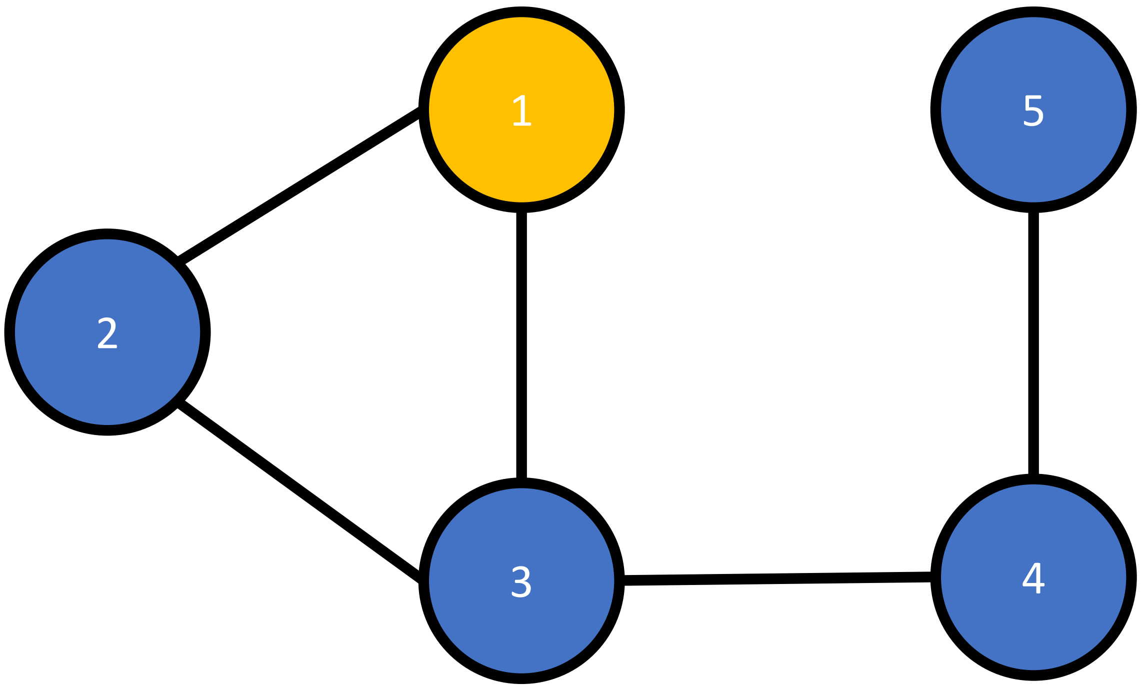
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



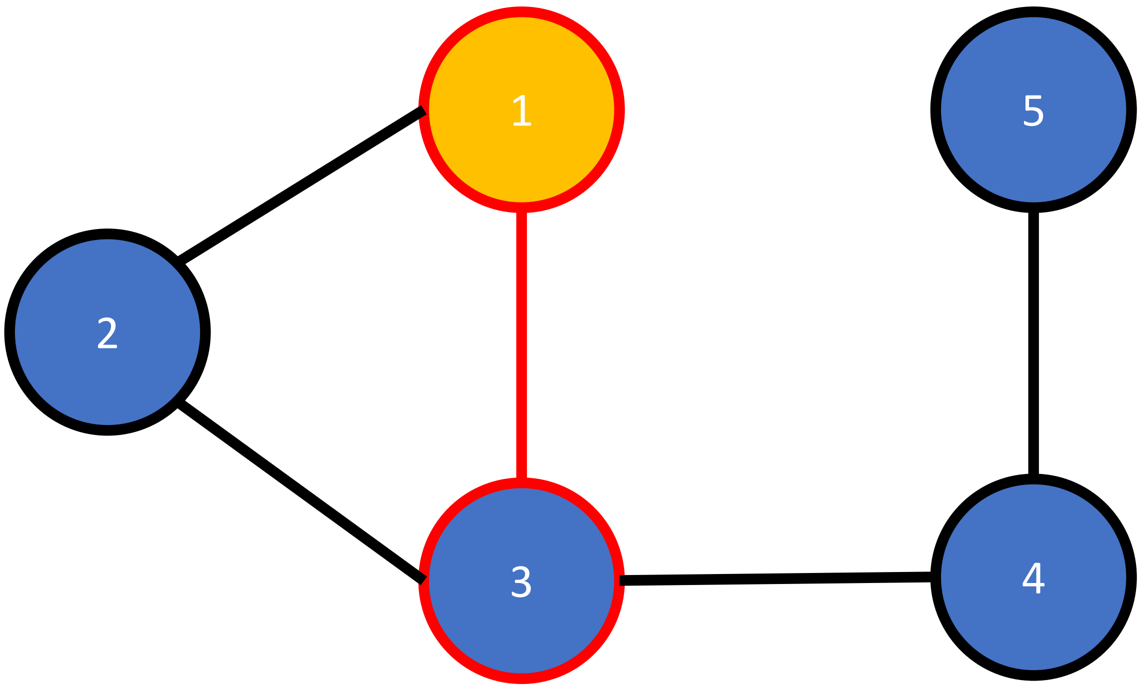
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



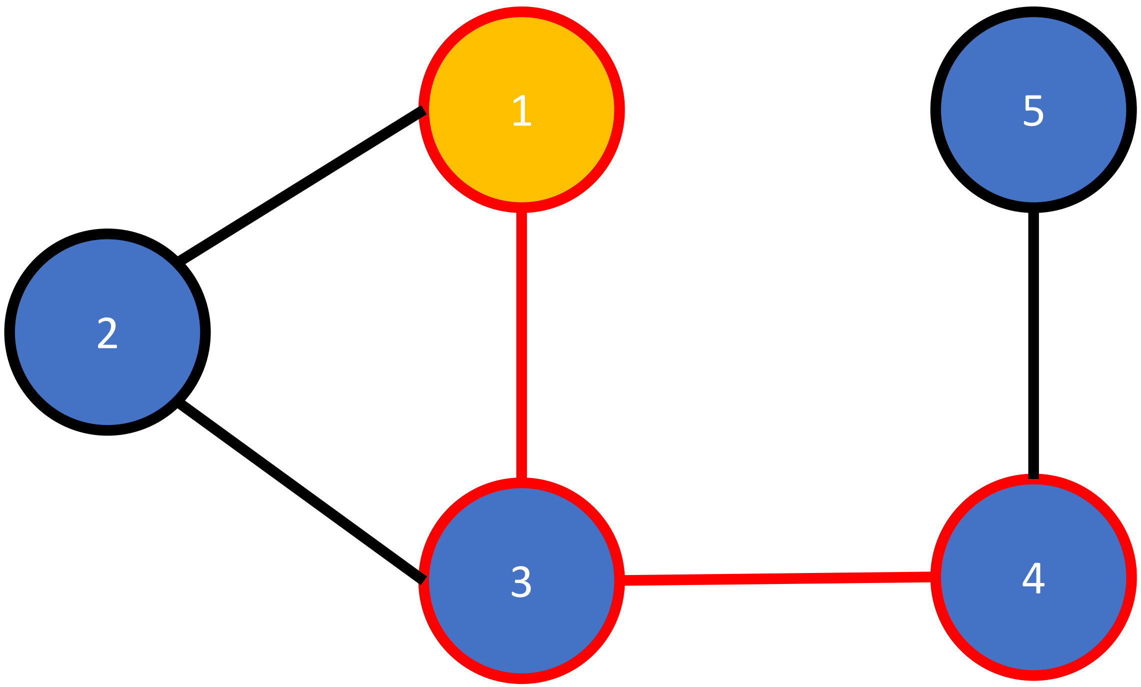
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



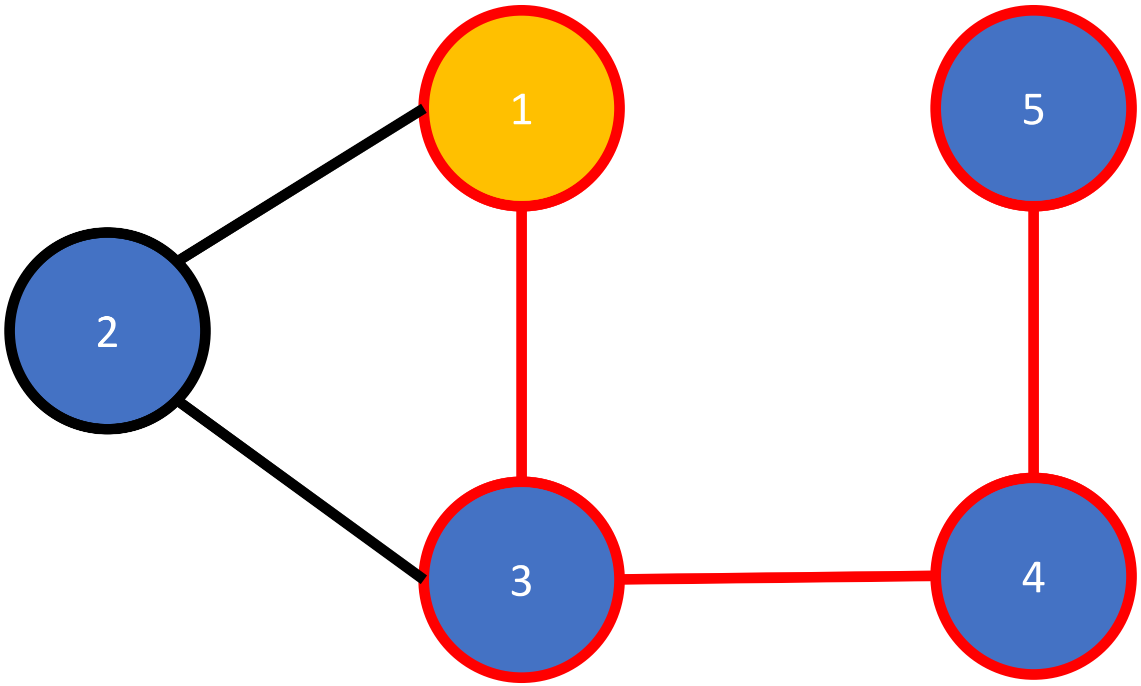
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



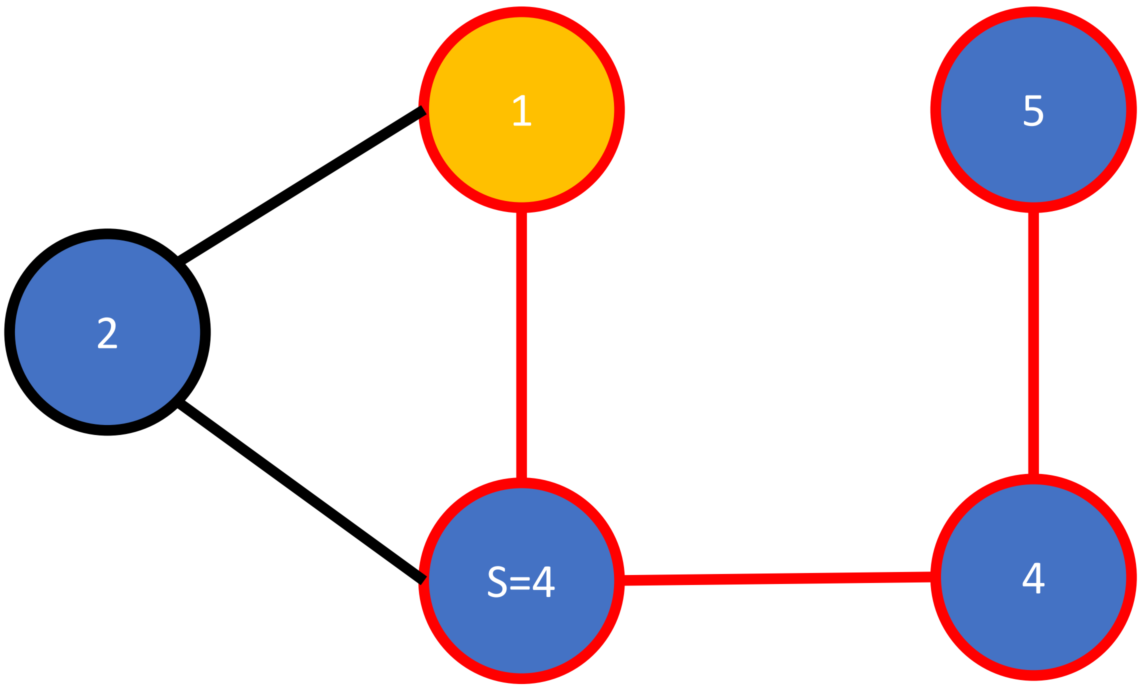
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



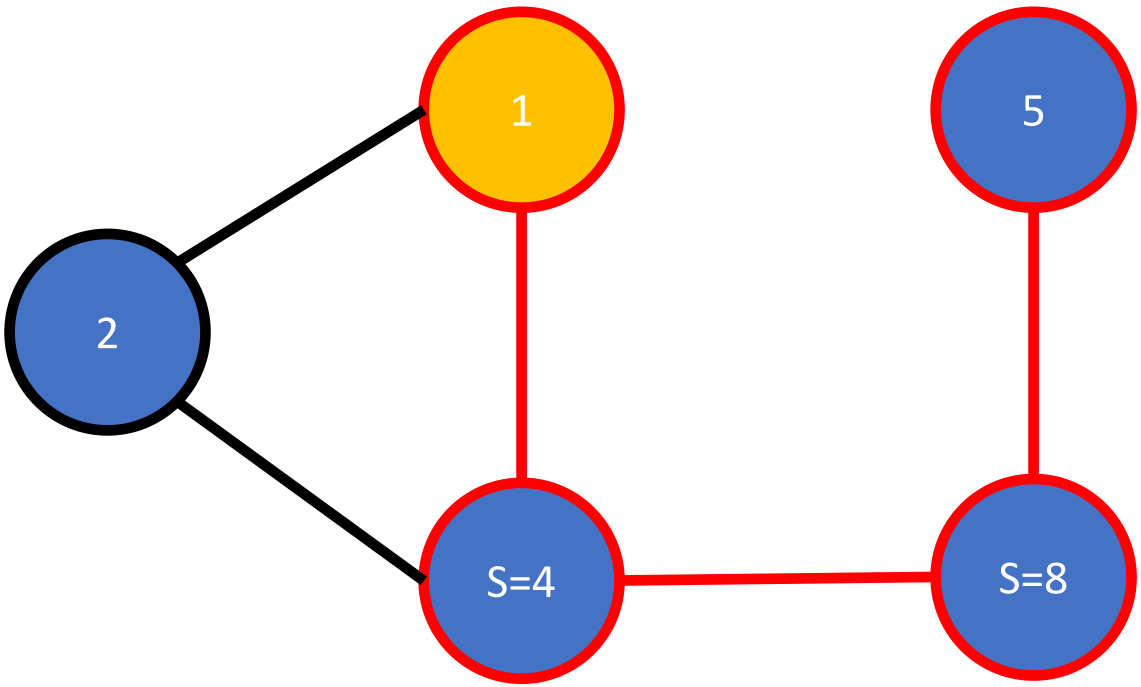
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



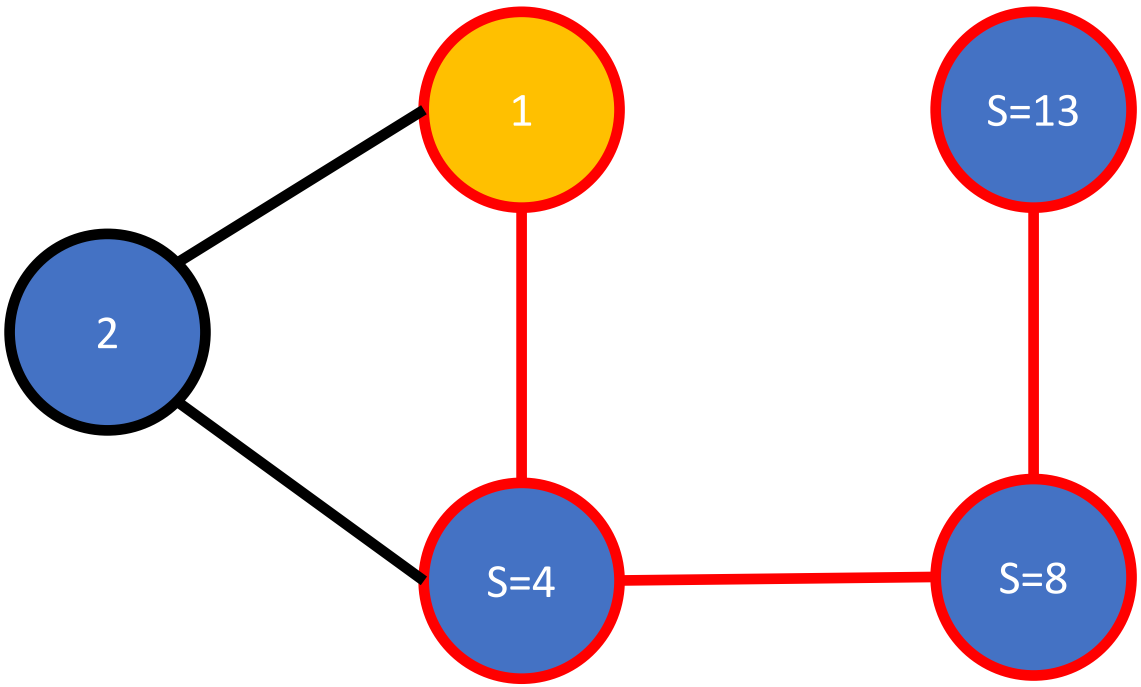
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



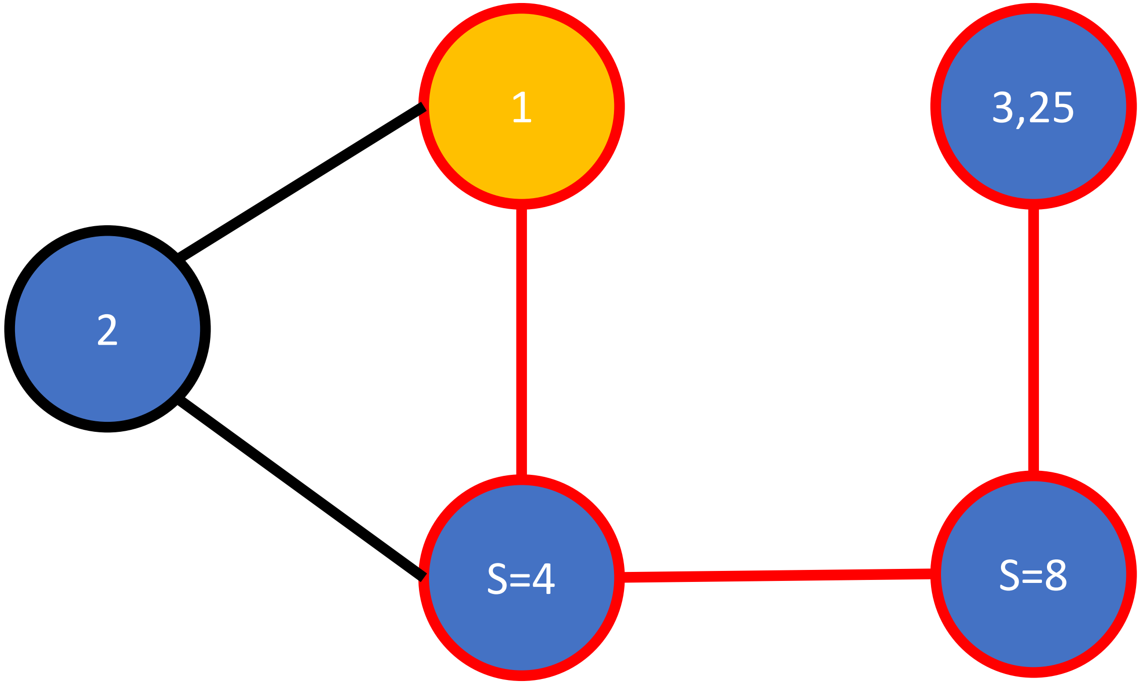
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



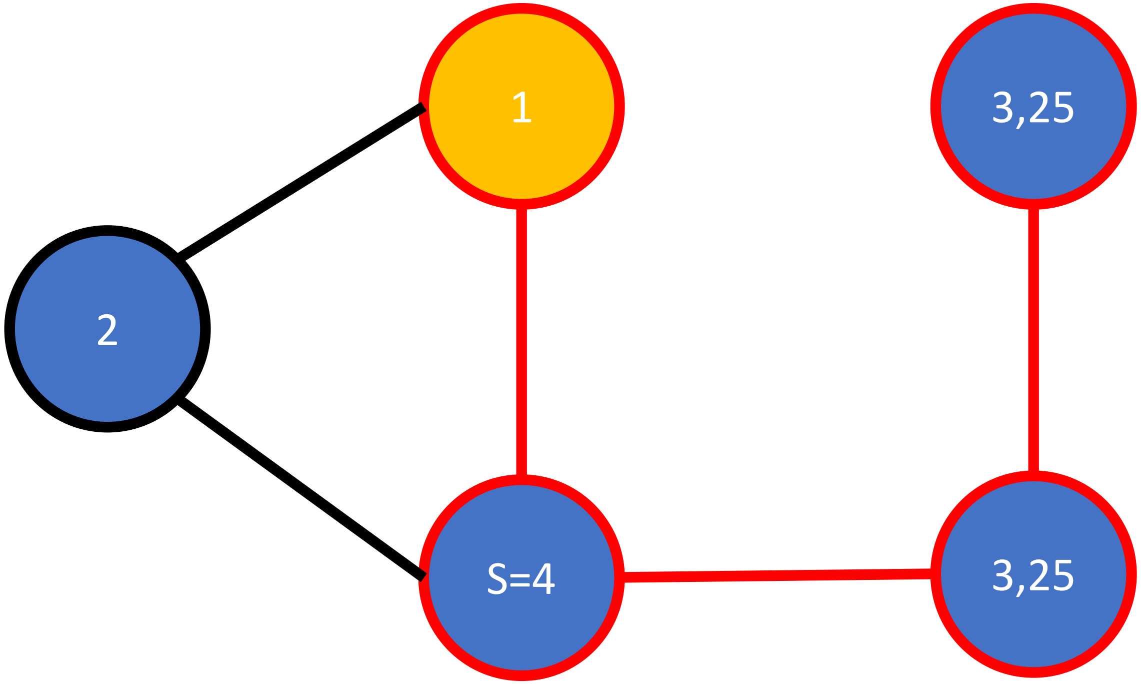
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



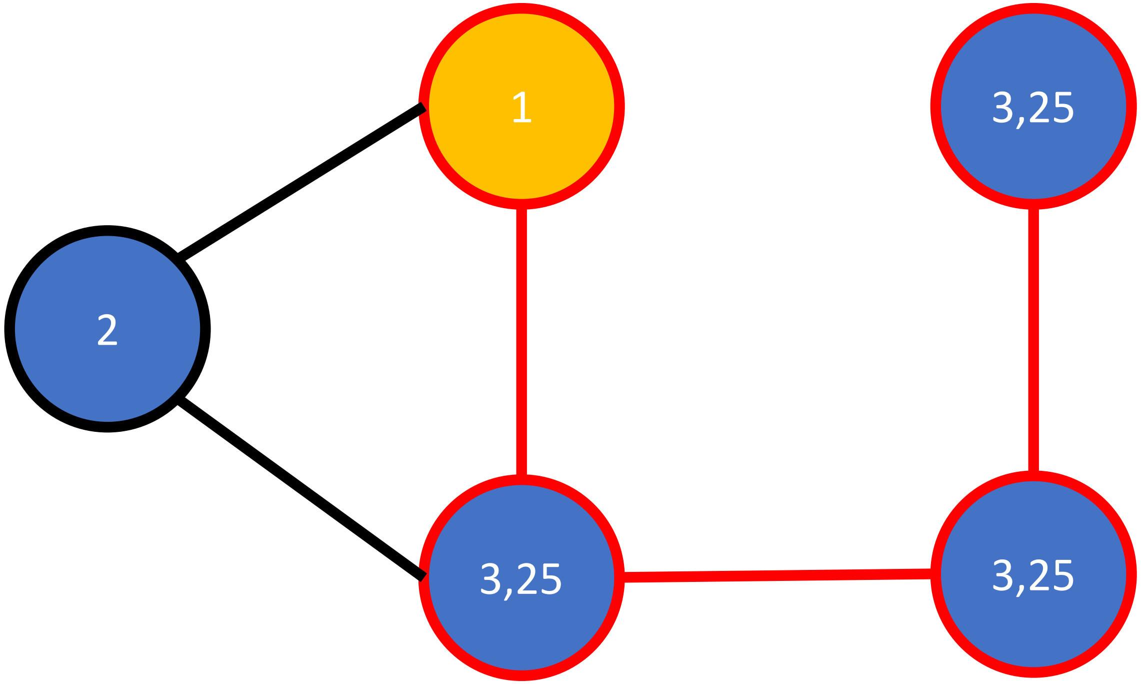
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



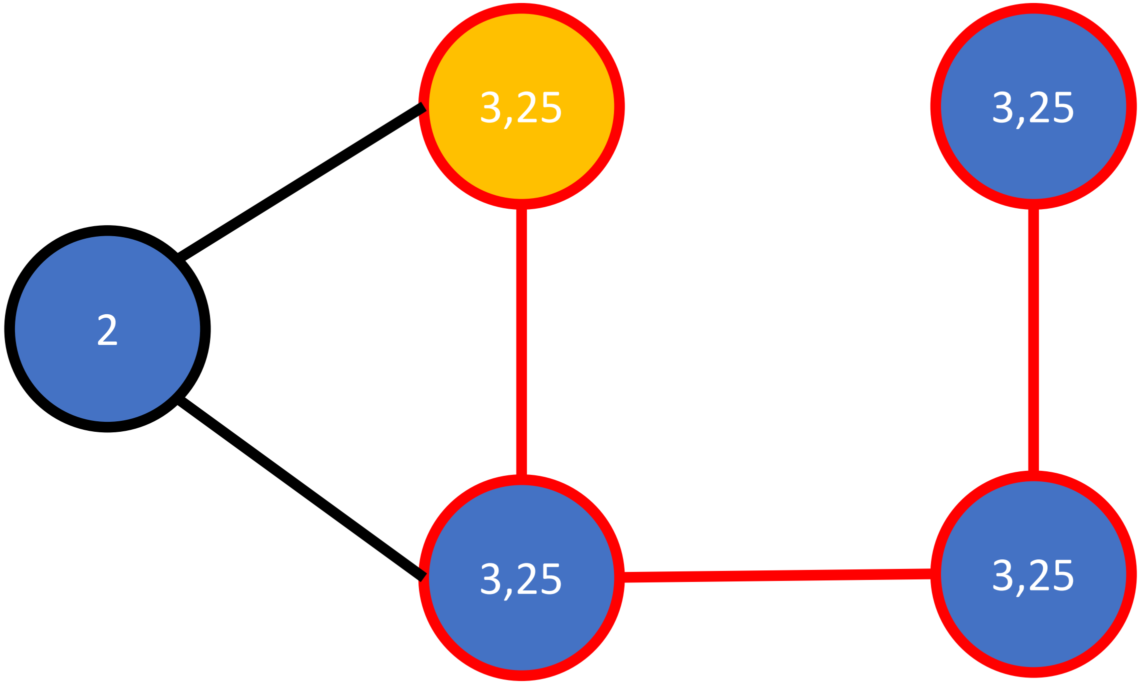
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



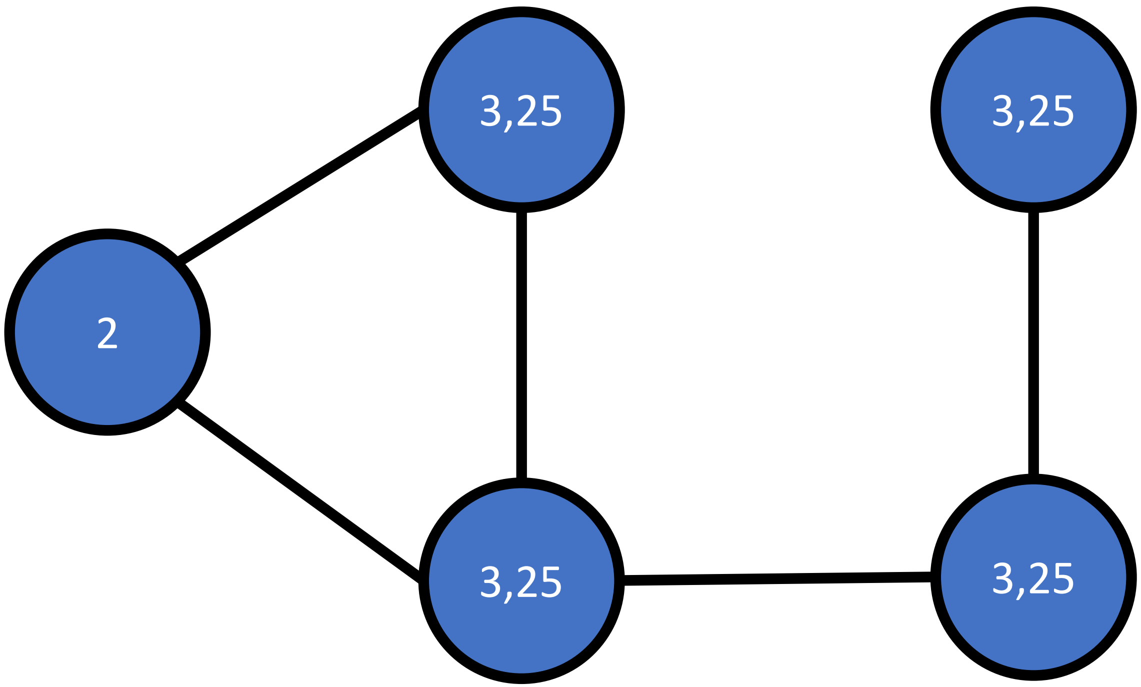
Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



Ο Αλγόριθμος Path Averaging

- Παράδειγμα: Ο μέσος όρος είναι (1+2+3+4+5)/5=15/5=3



Γενικοί Κανόνες Ανανέωσης

Γενικοί Κανόνες Ανανέωσης

- Μπορούμε να θεωρήσουμε έναν γενικό κανόνα με βάση τον οποίο ο κόμβος i ανανεώνει την εκτίμησή του x_i^t στη νέα τιμή x_i^{t+1}

$$\mathrm{x}_i^{t+1} = p_{ii}\mathrm{x}_i^t + \sum_{j\in n(i)} p_{ij}\mathrm{x}_j^t$$

όπου $n(i)$ το σύνολο των γειτονικών κόμβων του κόμβου i

- Οι τιμές των παραμέτρων p_{ij} καθορίζουν τις λεπτομέρειες του κανόνα ανανέωσης
- Μπορούμε να γράψουμε την πιο πάνω σχέση σε μορφή πινάκων

$$\boldsymbol{x}^{t+1} = \boldsymbol{P}\boldsymbol{x}^t$$

- $\boldsymbol{x}^t = [\mathrm{x}_1^t \quad \cdots \quad \mathrm{x}_n^t]^\mathrm{T}$
- $\boldsymbol{P} = \left[p_{ij}\right] \in \mathbb{R}^{n\times n}$
- και αρχικοποιούμε ως $\boldsymbol{x}^\mathbf{0} = [c_1 \quad \cdots \quad c_n]^\mathrm{T}$

Γενικοί Κανόνες Ανανέωσης

- Ορίζουμε το διάνυσμα που περιέχει σε όλες τις θέσεις του τη μέση τιμή $\bar{c}$ ως

$$\boldsymbol{x}_{ave} = \bar{c} \cdot \mathbf{1} = \left(\frac{1}{n}\mathbf{1}^{\mathrm{T}}\boldsymbol{x}^{\mathbf{0}}\right) \cdot \mathbf{1}$$

- Επίσης, ορίζουμε το διάνυσμα που περιέχει τα «σφάλματα» ως

$$\boldsymbol{y}^t = \boldsymbol{x}^t - \boldsymbol{x}_{ave}$$

- Έτσι, ο στόχος της συμφωνίας ως προς το μέσο όρο είναι να βεβαιωθούμε πως $\boldsymbol{y}^t \to \mathbf{0}$ καθώς $t \to \infty$, μέσω του κανόνα ανανέωσης

$$\boldsymbol{x}^{t+1} = \boldsymbol{P}\boldsymbol{x}^t$$

- Θεωρούμε τις ιδιοτιμές του πίνακα $\boldsymbol{P}$ ως εξής

$$\lambda_1(\boldsymbol{P}) \geq \lambda_2(\boldsymbol{P}) \geq \cdots \lambda_n(\boldsymbol{P})$$

Γενικοί Κανόνες Ανανέωσης

Θεώρημα

- Οι ακόλουθες τρείς συνθήκες είναι **αναγκαίες και ικανές** ώστε ο επαναληπτικός κανόνας

$$\boldsymbol{x}^{t+1} = \boldsymbol{P}\boldsymbol{x}^{t}$$

να πετύχει συμφωνία ως προς το μέσο όρο ξεκινώντας από οποιαδήποτε αρχική τιμή $\boldsymbol{x}^{\mathbf{0}}$

1. $\lambda_1(\boldsymbol{P}) = 1$ και επίσης $|\lambda_i(\boldsymbol{P})| < 1$ για όλα τα $i = 2, 3, \dots, n$
2. $\boldsymbol{P} \cdot \mathbf{1} = \mathbf{1}$, δηλαδή, το διάνυσμα $\mathbf{1}$ είναι ιδιοδιάνυσμα του πίνακα $\boldsymbol{P}$ και αντιστοιχεί στην ιδιοτιμή $\lambda_1(\boldsymbol{P}) = 1$
3. $\mathbf{1}^{\mathrm{T}} \cdot \boldsymbol{P} = \mathbf{1}^{\mathrm{T}}$, δηλαδή, το διάνυσμα $\mathbf{1}$ είναι και αριστερό ιδιοδιάνυσμα του πίνακα $\boldsymbol{P}$

Σημείωση: Οι συνθήκες 2 και 3 υπονοούν πως ο πίνακας $\boldsymbol{P}$ είναι doubly stochastic. Επίσης, η συνθήκη 2 καθορίζει πως αυτό που υπολογίζουμε είναι ο μέσος όρος, ενώ η συνθήκη 3 είναι απαραίτητη προκειμένου να έχουμε συμφωνία

Average Consensus ως Κατανεμημένη Βελτιστοποίηση

Average Consensus ως Βελτιστοποίηση

- Είναι αρκετά εύκολο να δούμε πως το πρόβλημα

$$\min_{x_1, x_2, \cdots, x_n} \left(\sum_{i=1}^{n} \frac{1}{2} (x_i - c_i)^2 \right) \quad s.t. \quad x_1 = x_2 = \cdots = x_n$$

- Έχει ως βέλτιστη λύση την

$$\boldsymbol{x}^* = \bar{c} \cdot \mathbf{1}$$

δηλαδή, έχει τη συνηθισμένη μορφή που έχουν τα προβλήματα τα οποία επιδέχονται κατανεμημένη βελτιστοποίηση

Ερώτηση: Ποια τεχνική θα διαλέγατε, και γιατί ?

- Την τεχνική χαλάρωσης στον «πρωτεύοντα χώρο»
- Την τεχνική dual ascend

Κατευθυνόμενα Γραφήματα

Αλγόριθμος Uniform Gossip

Συναίνεση σε Κατευθυνόμενα Γραφήματα

- Όταν το δίκτυό μας περιγράφεται από ένα κατευθυνόμενο γράφημα, δεν μπορούμε να βρούμε έναν doubly stochastic πίνακα $\boldsymbol{P}$
- Για να ξεπεράσουμε αυτό το πρόβλημα, κάθε κόμβος διατηρεί ένα άθροισμα s_i^t και ένα βάρος w_i^t
- Ορίζουμε τα διανύσματα $\boldsymbol{s}^t = [s_1^t \quad \cdots \quad \mathrm{s}_n^t]^{\mathrm{T}}$ και $\boldsymbol{w}^t = [w_1^t \quad \cdots \quad \mathrm{w}_n^t]^{\mathrm{T}}$

Uniform Gossip

- Αρχικοποιούμε $\boldsymbol{s}^{\mathbf{0}} = \bar{c} \cdot \mathbf{1}$ και $\boldsymbol{w}^{\mathbf{0}} = \mathbf{1}$
- Ανά πάσα στιγμή, το διάνυσμα εκτιμήσεων δίνεται ως $\boldsymbol{x}^t = \boldsymbol{s}^t$ ./ $\boldsymbol{w}^t$
- Στην επανάληψη t θεωρούμε έναν πίνακα $\boldsymbol{D}(t)$ όπου για κάθε κόμβο i έχουμε $\boldsymbol{D}_{ii}(t) = 1/2$ και επιλέγουμε έναν τυχαίο κόμβο j και θέτουμε $\boldsymbol{D}_{ij}(t) = 1/2$
- Ανανεώνουμε ως

$$\boldsymbol{s}^{t+1} = \boldsymbol{D}(t) \cdot \boldsymbol{s}^t$$

$$\boldsymbol{w}^{t+1} = \boldsymbol{D}(t) \cdot \boldsymbol{w}^t$$

Συναίνεση σε Κατευθυνόμενα Γραφήματα

- Πιο πρόσφατες εργασίες έχουν προτείνει πιο γενικά σχήματα
- Το πρόβλημα της βελτιστοποίησης σε κατευθυνόμενα γραφήματα είναι αυτή τη στιγμή ανοικτό ερευνητικά

Ερωτήσεις