

# Influence Maximization σε Social Networks

Βαγγέλης Χατζηαφράτης

28 Απριλίου 2015

## ΣΥΝΟΠΤΙΚΑ:

- ▶ Αν θέλω να πετύχω υιοθέτηση μιας άποψης/προϊόντος από πολλούς, πως διαλέγω το αρχικό target group (free samples)

## ΣΥΝΟΠΤΙΚΑ:

- ▶ Αν θέλω να πετύχω υιοθέτηση μιας άποψης/προϊόντος από πολλούς, πως διαλέγω το αρχικό target group (free samples)
- ▶ Πρόβλημα Βελτιστοποίησης: Βρες την  $k$ -άδα ανθρώπων με μεγάλη επιρροή (;)

## ΣΥΝΟΠΤΙΚΑ:

- ▶ Αν θέλω να πετύχω υιοθέτηση μιας άποψης/προϊόντος από πολλούς, πως διαλέγω το αρχικό target group (free samples)
- ▶ Πρόβλημα Βελτιστοποίησης: Βρες την  $k$ -άδα ανθρώπων με μεγάλη επιρροή (;)
- ▶ Επειδή το πρόβλημα βελτιστοποίησης είναι NP-hard σε όλα τα ενδιαφέροντα μοντέλα διάδοσης πληροφορίας θα δούμε έναν Greedy αλγόριθμο με εγγύηση προσέγγισης  $63\% (\approx 1 - 1/e)$  που προκύπτει από ένα βασικό θεώρημα στις submodular functions (;)

# Influence Maximization

- ▶ Ορίζω ως influence ενός συνόλου  $A$  το  $\sigma(A) =$  expected number of active nodes at the end of the process

# Influence Maximization

- ▶ Ορίζω ως influence ενός συνόλου  $A$  το  $\sigma(A) =$  expected number of active nodes at the end of the process
- ▶ Ορίζω ως influence maximization problem σε ένα γράφο, δοσμένου του  $k$  , να βρω  $k$ -άδα κόμβων με maximum influence

# Influence Maximization

- ▶ Ορίζω ως influence ενός συνόλου  $A$  το  $\sigma(A) =$  expected number of active nodes at the end of the process
- ▶ Ορίζω ως influence maximization problem σε ένα γράφο, δοσμένου του  $k$ , να βρω  $k$ -άδα κόμβων με maximum influence
- ▶ Στα ενδιαφέροντα μοντέλα LTM, ICM είναι NP-hard  
→ (*greedy 63% approx*)

## Submodular Functions

- ▶ natural diminishing returns property: το οριακό κέρδος του να προσθέσω ένα στοιχείο στο  $S$  είναι τουλάχιστον τόσο όσο το οριακό κέρδος του να προσθέσω το ίδιο στοιχείο σε ένα υπερσύνολο του  $S$  - σας θυμίζει κάτι αυτό από Algo ; (- άλλο 5 άλλο 1000)



# Submodular Functions

- ▶ natural diminishing returns property: το οριακό κέρδος του να προσθέσω ένα στοιχείο στο  $S$  είναι τουλάχιστον τόσο όσο το οριακό κέρδος του να προσθέσω το ίδιο στοιχείο σε ένα υπερσύνολο του  $S$  - σας θυμίζει κάτι αυτό από Algo ; (- άλλο 5 άλλο 1000)
- ▶ Με δεδομένη μία μονότονη, submodular μη αρνητική συνάρτηση, είναι NP-hard το να τη μεγιστοποιήσω σε σύνολα με σταθερό πληθάριθμο

## Greedy Algorithm - Εμπόδιο

- ▶ Ό,τι πιο φυσικό. Διάλεγε το κόμβο που σου μεγιστοποιεί το οριακό κέρδος. Πώς αφού έχω πιθανοτικά μοντέλα;

## Greedy Algorithm - Εμπόδιο

- ▶ Ό,τι πιο φυσικό. Διάλεγε το κόμβο που σου μεγιστοποιεί το οριακό κέρδος. Πώς αφού έχω πιθανοτικά μοντέλα;
- ▶ Πάμε με simulations/sampling γιατί υπάρχει θεώρημα που εγγυάται ότι οδηγούμαστε σε αυθαίρετα καλές προσεγγίσεις του  $\sigma(A)$  with high probability

# Greedy Algorithm - Εμπόδιο

- ▶ Ό,τι πιο φυσικό. Διάλεγε το κόμβο που σου μεγιστοποιεί το οριακό κέρδος. Πώς αφού έχω πιθανοτικά μοντέλα;
- ▶ Πάμε με simulations/sampling γιατί υπάρχει θεώρημα που εγγυάται ότι οδηγούμαστε σε αυθαίρετα καλές προσεγγίσεις του  $\sigma(A)$  with high probability
- ▶  $\forall \epsilon > 0, \exists \gamma > 0$  so that  $(1 + \gamma)$  stepwise approximations  $\rightarrow$   $(1 - 1/e - \epsilon)$  approximation of the final value

## Linear Threshold Model

- ▶ Βασίζεται στη λογική ότι ένας κόμβος/άνθρωπος τείνει όλο και περισσότερο να ενεργοποιηθεί/πειστεί/μολυνθεί όσο περισσότεροι φίλοι του είναι active/informed/infected

## Linear Threshold Model

- ▶ Βασίζεται στη λογική ότι ένας κόμβος/άνθρωπος τείνει όλο και περισσότερο να ενεργοποιηθεί/πειστεί/μολυνθεί όσο περισσότεροι φίλοι του είναι active/informed/infected
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active

# Linear Threshold Model

- ▶ Βασίζεται στη λογική ότι ένας κόμβος/άνθρωπος τείνει όλο και περισσότερο να ενεργοποιηθεί/πειστεί/μολυνθεί όσο περισσότεροι φίλοι του είναι active/informed/infected
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active
- ▶ Κάθε κόμβος  $v$  διαλέγει ομοιόμορφα και τυχαία ένα κατώφλι  $\theta_v \in [0, 1]$

# Linear Threshold Model

- ▶ Βασίζεται στη λογική ότι ένας κόμβος/άνθρωπος τείνει όλο και περισσότερο να ενεργοποιηθεί/πειστεί/μολυνθεί όσο περισσότεροι φίλοι του είναι active/informed/infected
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active
- ▶ Κάθε κόμβος  $v$  διαλέγει ομοιόμορφα και τυχαία ένα κατώφλι  $\theta_v \in [0, 1]$
- ▶ Πείθεται αν

$$\sum_{w \in N(v)} b_{v,w} \geq \theta_v, \text{ given } \sum_{w \in N(v)} b_{v,w} \leq 1, \forall v \in V$$



# Independent Cascade Model

- ▶ Βασίζεται στη λογική ότι ένας active κόμβος/άνθρωπος μπορεί να ενεργοποιήσει/πείσει/μολύνει κάποιους από τους φίλους του

# Independent Cascade Model

- ▶ Βασίζεται στη λογική ότι ένας active κόμβος/άνθρωπος μπορεί να ενεργοποιήσει/πείσει/μολύνει κάποιους από τους φίλους του
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active

# Independent Cascade Model

- ▶ Βασίζεται στη λογική ότι ένας active κόμβος/άνθρωπος μπορεί να ενεργοποιήσει/πείσει/μολύνει κάποιους από τους φίλους του
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active
- ▶ Κάθε κόμβος  $v$  προσπαθεί άπαξ να πείσει τους γείτονές του και πετυχαίνει (ανεξάρτητα) με πιθανότητα  $p_{v,w}$  όπου  $w$  γείτονας του  $v$

# Independent Cascade Model

- ▶ Βασίζεται στη λογική ότι ένας active κόμβος/άνθρωπος μπορεί να ενεργοποιήσει/πείσει/μολύνει κάποιους από τους φίλους του
- ▶ Progressive γιατί κάθε κόμβος από inactive περνά σε active
- ▶ Κάθε κόμβος  $v$  προσπαθεί άπαξ να πείσει τους γείτονές του και πετυχαίνει (ανεξάρτητα) με πιθανότητα  $p_{v,w}$  όπου  $w$  γείτονας του  $v$
- ▶ Αν τύχει και ενεργοποιηθούν ταυτόχρονα πολλοί γείτονες του  $v$  τότε οι προσπάθειες με αυθαίρετη (; - κέρματα reachability) σειρά

## Reachability, κέρματα και ICM submodularity

- ▶ Ισοδύναμα στο ICM εξαρχής ρίχνω κέρματα για κάθε ακμή με επιτυχία ανάλογη της πιθανότητάς της

## Reachability, κέρματα και ICM submodularity

- ▶ Ισοδύναμα στο ICM εξαρχής ρίχνω κέρματα για κάθε ακμή με επιτυχία ανάλογη της πιθανότητάς της
- ▶ Δημιουργώ τον γράφο με ζωντανές ακμές μόνον και κάνω BFS από το A

## Reachability, κέρματα και ICM submodularity

- ▶ Ισοδύναμα στο ICM εξαρχής ρίχνω κέρματα για κάθε ακμή με επιτυχία ανάλογη της πιθανότητάς της
- ▶ Δημιουργώ τον γράφο με ζωντανές ακμές μόνον και κάνω BFS από το A
- ▶ Έχοντας fixάρι την κατανομή των κερμάτων ποια είναι η  $\sigma(A)$  ; Είναι submodular ;

## ICM Αναγωγή - Σχόλια για LTM

- ▶ To Influence Maximization - ICM είναι NP-Hard



## ICM Αναγωγή - Σχόλια για LTM

- ▶ Το Influence Maximization - ICM είναι NP-Hard
- ▶ Αναγωγή σε Κάλυμμα Συνόλου

## ICM Αναγωγή - Σχόλια για LTM

- ▶ Το Influence Maximization - ICM είναι NP-Hard
- ▶ Αναγωγή σε Κάλυμμα Συνόλου
- ▶ Και για το LTM ισχύουν τα ίδια συμπεράσματα.

## GTM , GCM

- ▶ Αφήνουμε τη συνάρτηση κατωφλίου να εξαρτάται από τα σύνολα γειτόνων.

## GTM , GCM

- ▶ Αφήνουμε τη συνάρτηση κατωφλίου να εξαρτάται από τα σύνολα γειτόνων.
- ▶ Επιτρέπουμε η πιθανότητα επιτυχίας να εξαρτάται από τους γείτονες που προσπάθησαν και απέτυχαν ήδη.

## GTM , GCM

- ▶ Αφήνουμε τη συνάρτηση κατωφλίου να εξαρτάται από τα σύνολα γειτόνων.
- ▶ Επιτρέπουμε η πιθανότητα επιτυχίας να εξαρτάται από τους γείτονες που προσπάθησαν και απέτυχαν ήδη.
- ▶ Τα παραπάνω αποδεικνύονται ισοδύναμα.

## GTM , GCM

- ▶ Αφήνουμε τη συνάρτηση κατωφλίου να εξαρτάται από τα σύνολα γειτόνων.
- ▶ Επιτρέπουμε η πιθανότητα επιτυχίας να εξαρτάται από τους γείτονες που προσπάθησαν και απέτυχαν ήδη.
- ▶ Τα παραπάνω αποδεικνύονται ισοδύναμα.
- ▶ Στα γενικά αυτά μοντέλα η προσέγγιση σε παράγοντα  $n^{1-\epsilon}$  του Influence Maximization - ICM είναι NP-Hard (;

## Triggering Models & Conjecture

- ▶ Κάθε κορυφή διαλέγει ανεξάρτητα με βάση κάποια κατανομή πιθανότητας στα υποσύνολα των γειτόνων της ένα τυχαίο Triggering Set  $T_v$  και περνά από inactive to active αν έχει γείτονα στο Triggering Set  $T_v$  που ήταν active στο προηγούμενο βήμα

## Triggering Models & Conjecture

- ▶ Κάθε κορυφή διαλέγει ανεξάρτητα με βάση κάποια κατανομή πιθανότητας στα υποσύνολα των γειτόνων της ένα τυχαίο Triggering Set  $T_v$  και περνά από inactive to active αν έχει γείτονα στο Triggering Set  $T_v$  που ήταν active στο προηγούμενο βήμα
- ▶ Σε αυτό το λίγο πιο ειδικό μοντέλο πάλι έχω submodularity & greedy



## Triggering Models & Conjecture

- ▶ Κάθε κορυφή διαλέγει ανεξάρτητα με βάση κάποια κατανομή πιθανότητας στα υποσύνολα των γειτόνων της ένα τυχαίο Triggering Set  $T_v$  και περνά από inactive to active αν έχει γείτονα στο Triggering Set  $T_v$  που ήταν active στο προηγούμενο βήμα
- ▶ Σε αυτό το λίγο πιο ειδικό μοντέλο πάλι έχω submodularity & greedy
- ▶ Whenever the threshold functions  $f_v$  at every node are monotone and submodular, the resulting influence function is monotone and submodular as well

## Τι είδαμε και τι θα δούμε

- ▶ Διάφορα Μοντέλα διάδοσης πληροφορίας και σχέση με submodularity

## Τι είδαμε και τι θα δούμε

- ▶ Διάφορα Μοντέλα διάδοσης πληροφορίας και σχέση με submodularity
- ▶ Αν έχω πολλαπλά δίκτυα τί γίνεται ;

## Τι είδαμε και τι θα δούμε

- ▶ Διάφορα Μοντέλα διάδοσης πληροφορίας και σχέση με submodularity
- ▶ Αν έχω πολλαπλά δίκτυα τί γίνεται ;
- ▶ Σύνδεση με Game Theory/Opinion Dynamics και εταιρείες.

# Ευχαριστώ για την προσοχή σας!

► Απορίες