

Αναφορά

Μάθηση σε Συνθήκες Διαδικτύου για το Min Sum Set Cover και το Κουτί της Πανδώρας

Ιωάννης Κασιώνης
Πανεπιστήμιο Πειραιώς
ioannis.kasionis@gmail.com

Πέτρος Τριαντάφυλλος
Πανεπιστήμιο Πειραιώς
petrostriantafyllos@outlook.com

Ιανουάριος 2025

Abstract

Η παρούσα αναφορά συνοψίζει και αναλύει το άρθρο *Online Learning for Min Sum Set Cover and Pandora's Box* [1], το οποίο προτείνει έναν απλό αλλά αποτελεσματικό «σκελετό» για τη σχεδίαση αλγορίθμων μάθησης σε συνθήκες διαδικτύου (*online learning*) για το Πρόβλημα του Κουτιού της Πανδώρας (*Pandora's Box*), το Πρόβλημα Min Sum Set Cover (MSSC), καθώς και συναφή προβλήματα στο πεδίο της στοχαστικής βελτιστοποίησης (*Stochastic optimization*). Παραδοσιακά, τα συγκεκριμένα προβλήματα θεωρούν ότι οι κατανομές των τιμών είναι εκ των προτέρων γνωστές. Στο πλαίσιο όμως αυτής της εργασίας, εξετάζεται μια προσέγγιση μάθησης σε συνθήκες διαδικτύου, όπου σε κάθε γύρο αποκαλύπτεται νέος πληροφοριακός όγκος δεδομένων.

Επιπροσθέτως, το άρθρο επεκτείνει το πλαίσιο αυτό και σε ένα σενάριο *bandit*, στο οποίο μετά το τέλος κάθε γύρου ο αλγόριθμος μαθαίνει μόνο τις τιμές των κουτιών που αποφάσισε να «ανοίξει». Αυτή η επέκταση καλύπτει και εναλλακτικές μορφές των παραπάνω προβλημάτων [1].

1 Εισαγωγή

Τα δύο βασικά προβλήματα στη *στοχαστική βελτιστοποίηση* είναι το *Κουτί της Πανδώρας* (*Pandora's Box*) και το *Min Sum Set Cover (MSSC)*. Το *Κουτί της Πανδώρας*, που εισήχθη από τον Weitzman [2], αποτελεί ένα γενικό πλαίσιο λήψης αποφάσεων υπό αβεβαιότητα, με στόχο την ελαχιστοποίηση του κόστους που σχετίζεται με τη συλλογή πληροφοριών. Το *MSSC*, ως παραλλαγή του *Κουτιού της Πανδώρας*, επικεντρώνεται σε περιπτώσεις όπου οι τιμές είναι είτε 0 είτε ∞ [3].

Ενώ η κλασική διατύπωση των συγκεκριμένων προβλημάτων υποθέτει γνωστή κατανομή για τις τιμές και τα κόστη [1], η παρούσα εργασία επεκτείνει το πλαίσιο σε *online* συνθήκες, όπου τα δεδομένα καταφθάνουν σταδιακά, χωρίς να υπάρχει εκ των προτέρων γνώση για τις κατανομές.

Αυτή η *online learning* προσέγγιση εισάγει προκλήσεις αλλά και ευκαιρίες. Σε κάθε βήμα, είτε τα σενάρια είναι ανταγωνιστικά (adversarial) είτε ανεξάρτητα, το σύστημα καλείται να λάβει απόφαση βασισμένη μόνο στην τρέχουσα κατάσταση και σε προηγούμενες παρατηρήσεις. Η αποδοτικότητα συχνά μετριέται με βάση τη *μετανάστευση* (*regret*), η οποία συγκρίνει το σωρευτικό κόστος του αλγορίθμου με αυτό μιας βέλτιστης εκ των υστέρων στρατηγικής [4].

Κύριες συνεισφορές

1. Παρουσίαση ενός αποδοτικού ως προς τον υπολογισμό *online* αλγορίθμου, που επιτυγχάνει σταθερή ανταγωνιστικότητα (constant-competitive) για διατυπώσεις τόσο του *MSSC* όσο και του Προβλήματος του Κουτιού της Πανδώρας, σε συνθήκες πλήρους πληροφορίας. Αυτό περιλαμβάνει προβλήματα με πιο σύνθετους περιορισμούς, όπως βάσεις ματρεϊδών [5].
2. Επέκταση της ανάλυσης σε *bandit* συνθήκες, όπου υπάρχει περιορισμένη πληροφόρηση σχετικά με τις τιμές που αποκαλύπτονται [6]. Παρότι οι παρατηρήσεις είναι ελλιπείς, ο προτεινόμενος αλγόριθμος διατηρεί λογαριθμική μετανάστευση, διασφαλίζοντας υψηλή απόδοση.
3. Αξιοποίηση κυρτών χαλαρώσεων (convex relaxations) και τεχνικών *online* κυρτής βελτιστοποίησης, για να επιτευχθούν πρακτικές προσεγγίσεις (approximations) που γεφυρώνουν τη θεωρία με την αποδοτική υλοποίηση [7].

Οι συνέπειες αυτών των αποτελεσμάτων είναι διττές. Πρώτον, επεκτείνουν την εφαρμοσιμότητα του *MSSC* και του Κουτιού της Πανδώρας σε δυναμικά πλαίσια του πραγματικού κόσμου, όπου οι αποφάσεις πρέπει να προσαρμόζονται σε εξελισσόμενα δεδομένα. Δεύτερον, η ενσωμάτωση *online* τεχνικών κυρτής βελτιστοποίησης υπογραμμίζει τη σημασία μαθηματικών εργαλείων για την αντιμετώπιση συνδυαστικών προβλημάτων [1].

Η υπόλοιπη αναφορά δομείται ως εξής: Στην Ενότητα 2 παρουσιάζονται οι διατυπώσεις των προβλημάτων και η σχετική βιβλιογραφία. Στην Ενότητα 3 περιγράφονται οι προτεινόμενες μέθοδοι. Οι Ενότητες 4 και 5 περιλαμβάνουν την ανάλυση και τα πειραματικά αποτελέσματα αντίστοιχα. Τέλος, παρουσιάζονται συζητήσεις για πιθανές μελλοντικές επεκτάσεις.

2 Ορισμοί Προβλημάτων

2.1 Πρόβλημα Κουτιού της Πανδώρας (Pandora's Box Problem)

Στο Πρόβλημα του Κουτιού της Πανδώρας, δίνεται ένα σύνολο από n κουτιά, το καθένα με άγνωστη τιμή και ένα κόστος για να ανοιχτεί. Σκοπός είναι η εύρεση της σειράς με την οποία θα ανοιχτούν τα κουτιά, ώστε να ελαχιστοποιηθεί το συνολικό κόστος της διερεύνησης και της τελικής επιλογής. Πιο συγκεκριμένα, δεδομένης μίας κατανομής πιθανών τιμών και εξόδων για κάθε κουτί, επιδιώκεται μία στρατηγική που ισορροπεί μεταξύ της εξερεύνησης (άνοιγμα κουτιών) και της εκμετάλλευσης (επιλογή της βέλτιστης τιμής).

Στη διαδικτυακή (online) παραλλαγή του προβλήματος που μελετάται στην εργασία, οι τιμές και τα κόστη των κουτιών αποκαλύπτονται ανταγωνιστικά (adversarially) σε T γύρους, αναγκάζοντας το σύστημα να προσαρμόζεται διαρκώς στις νέες πληροφορίες, διατηρώντας χαμηλή μετανάστευση (regret) σε σύγκριση με μια βέλτιστη εκ των υστέρων πολιτική.

2.2 Min Sum Set Cover (MSSC)

Το MSSC συνιστά ειδική περίπτωση του Προβλήματος του Κουτιού της Πανδώρας, όπου οι τιμές (ή τα κόστη) είναι είτε 0 είτε ∞ . Το MSSC εφαρμόζεται σε περιπτώσεις όπου ένα στοιχείο είτε ικανοποιεί απόλυτα κάποια κριτήρια είτε όχι (binary outcomes).

3 Μεθοδολογίες και Αποτελέσματα

Οι συγγραφείς προτείνουν ένα τριφασικό πλαίσιο, εμπνευσμένο από τις *scenario-aware* χαλαρώσεις του [?], προσαρμοσμένο όμως σε ένα *online* περιβάλλον. Το πλαίσιο εφαρμόζεται σε δύο κριτήρια σύγκρισης (*benchmarks*), το *non-adaptive (NA)* και το *partially adaptive (PA)*, τόσο σε συνθήκες πλήρους πληροφoρίας όσο και σε *bandit* ανατροφοδότηση.

Scenario-Aware Χαλάρωση και Στρογγυλοποίηση (Rounding)

Αρχικά, σε κάθε γύρο διατυπώνεται το εκάστοτε πρόβλημα (επιλογή ενός κουτιού, επιλογή k κουτιών, επιλογή βάσης ματρειδούς) ως ένα *scenario-aware* κυρτό πρόγραμμα. Η κλασματική λύση του προγράμματος αντιπροσωπεύει την «πιθανότητα» ανοίγματος ή επιλογής κουτιών. Βασισζόμενο στο [?], το σχήμα χαλάρωσης σχεδιάζεται έτσι ώστε μια ενιαία διαδικασία στρογγυλοποίησης (*rounding*) να μπορεί να εφαρμοστεί ανεξάρτητα από το σενάριο που πραγματικά αποκαλύπτεται. Ένα α -προσεγγιστικό (*approximate*) σχήμα στρογγυλοποίησης εγγυάται ότι η μετατροπή της κλασματικής λύσης σε ακέραιη (δηλ. πραγματική ακολουθία ή υποσύνολο κουτιών) αυξάνει το κόστος το πολύ κατά έναν σταθερό παράγοντα α . Έτσι, αποφεύγεται η ανάγκη επαναβέλτιστης λύσης σε κάθε διαφορετικό σενάριο και διατηρούνται οι ιδιότητες *no-regret*.

Online Κυρτή Βελτιστοποίηση σε Περιβάλλον Πλήρους Πληροφoρίας έναντι PA

Για την αντιμετώπιση ενός *partially adaptive* προτύπου (*benchmark*) σε περιβάλλον πλήρους πληροφoρίας, οι συγγραφείς χρησιμοποιούν τη μέθοδο *Follow the Regularized Leader (FTRL)*. Σε κάθε γύρο, αφού παρατηρηθεί η συνάρτηση κόστους, η λύση ανανεώνεται ελαχιστοποιώντας το άθροισμα παρελθουσών απωλειών συν έναν ισχυρά κυρτό κανονικοποιητή (*regularizer*). Αντί να εφαρμόζεται *gradient descent*, η μέθοδος FTRL προσφέρει πιο αυστηρά όρια στη μετανάστευση (*regret*). Στη συνέχεια, η κλασματική λύση στρογγυλοποιείται χρησιμοποιώντας το ενιαίο α -approximation σχήμα που προαναφέρθηκε. Στην ειδική περίπτωση του MSSC, το εν λόγω πλαίσιο ανακτά την ήδη γνωστή εκτός σύνδεσης (*offline*) 4-προσεγγιστική λύση, δείχνοντας ότι η προσέγγιση επιτυγχάνει τα βέλτιστα γνωστά αποτελέσματα ακόμα και σε ανταγωνιστικές ροές δεδομένων.

Επιπλέον, απέναντι σε ένα *partially adaptive* πρότυπο, οι μέθοδοι πετυχαίνουν α -προσεγγιστική *no-regret* σε στόχους τύπου Κουτιού της Πανδώρας, υπό διάφορους περιορισμούς. Συγκεκριμένα:

$$\begin{cases} 9.22\text{-προσεγγιστική no-regret για επιλογή 1 κουτιού,} \\ O(1)\text{-προσεγγιστική no-regret για επιλογή } k \text{ κουτιών,} \\ O(\log k)\text{-προσεγγιστική no-regret για επιλογή βάσης ματρειδούς.} \end{cases}$$

Τα αποτελέσματα αυτά ισχύουν σε καθεστώς πλήρους πληροφoρίας, επιτυγχάνοντας σταθερό ή λογαριθμικό πολλαπλασιαστικό παράγοντα έναντι μιας βέλτιστης εκ των υστέρων στρατηγικής *partially adaptive*.

Bandit έναντι PA μέσω μίξης FTRL και «Άνοιγμα Όλων»

Σε συνθήκες *bandit* ανατροφοδότησης, ο αλγόριθμος παρατηρεί μόνο τη συνάρτηση κόστους $f^{(t)}(x_t)$ στο σημείο x_t που επιλέγει. Κατά συνέπεια, δεν διαθέτει πλήρη γνώση της συνάρτησης σε κάθε γύρο. Για να αντιμετωπίσουν αυτόν τον περιορισμό, οι συγγραφείς προτείνουν μια τυχαία εναλλαγή μεταξύ:

1. Άνοιγμα όλων των n κουτιών, επωμιζόμενοι πλήρως το κόστος, ώστε να «μάθουν» τη συνάρτηση κόστους ολόκληρου του γύρου.
2. Εφαρμογή ενός βήματος OCO (όπως το FTRL), αξιοποιώντας τις περιορισμένες παρατηρήσεις από τα κουτιά που ανοίχτηκαν στους υπόλοιπους γύρους.

Αυτή η στρατηγική μίξης (exploration-exploitation) διατηρεί τον ίδιο α -προσεγγιστικό παράγοντα στοργγυλοποίησης και εξασφαλίζει υπογραμμική μετανάστευση (*sublinear regret*) σε σχέση με ένα PA πρότυπο. Η ισορροπία μεταξύ πλήρους εξερεύνησης (άνοιγμα όλων των κουτιών) και φειδωλής αξιοποίησης (επιλογή δειγματοληπτικών κουτιών) επιτυγχάνει σταθερή ή $O(\log k)$ προσεγγιστική απόδοση, παρόμοια με την περίπτωση πλήρους πληροφορίας.

Bandit έναντι NA και ο ρόλος των μη-Λιπσιτσιανών (non-Lipschitz) κόστων

Όταν το πρότυπο είναι *non-adaptive*, επιλέγεται ένα υποσύνολο κουτιών που παραμένει σταθερό σε όλους τους γύρους. Σε τέτοιες περιπτώσεις, οι συναρτήσεις κόστους μπορεί να είναι *non-Lipschitz*, καθιστώντας μη εφαρμόσιμη τη μέθοδο FTRL. Ως εκ τούτου, η εργασία παρουσιάζει μια διαδικασία *explore-exploit*:

1. **Εξερεύνηση (Explore):** Ο αλγόριθμος ανοίγει όλα τα κουτιά σε ορισμένους γύρους, ώστε να ανακαλύψει νέους περιορισμούς (π.χ. ένα όριο τιμής). Αυτές οι πληροφορίες ενσωματώνονται σε ένα παγκόσμιο γραμμικό πρόγραμμα (linear program) που επιλύεται με τη μέθοδο του ελλειψοειδούς (ellipsoid).
2. **Εκμετάλλευση (Exploit):** Η κλασματική λύση που προκύπτει στοργγυλοποιείται σε ένα μη-προσαρμόσιμο (*non-adaptive*) σύνολο κουτιών, το οποίο διατηρείται σταθερό μέχρι την επόμενη φάση εξερεύνησης.

Το Θεώρημα 5.3 δηλώνει ότι αν υπάρχει ένας *partially adaptive* αλγόριθμος με συντελεστή ανταγωνιστικότητας β σε σχέση με το βέλτιστο NA , τότε ο προτεινόμενος *bandit* αλγόριθμος φτάνει την ίδια β έως έναν μικρό σταθερό παράγοντα και ένα αμελητέο $o(T)$ όρο στη μετανάστευση. Χρησιμοποιώντας γνωστούς συντελεστές από προηγούμενα αποτελέσματα, προκύπτουν τα εξής όρια για *approximate no-regret*:

$$\begin{cases} 3.16\text{-προσεγγιστική no-regret για επιλογή 1 κουτιού,} \\ 12.64\text{-προσεγγιστική no-regret για επιλογή } k \text{ κουτιών,} \\ O(\log k)\text{-προσεγγιστική no-regret για επιλογή βάσης ματροειδούς.} \end{cases}$$

Άρα, ακόμη και στη *non-adaptive* συνθήκη και με μερική πληροφόρηση (*bandit*), το σχήμα διατηρεί τη δυνατότητα επίτευξης σχεδόν βέλτιστων εγγυήσεων ανταγωνιστικότητας.

Συμπεράσματα

Συνολικά, η μεθοδολογία αυτή — βασισμένη σε *scenario-aware* κυρτές χαλαρώσεις, *online* κυρτή βελτιστοποίηση και ανεξάρτητη από το σενάριο στοργγυλοποίηση — προσφέρει *σταθερής ανταγωνιστικότητας no-regret* αλγορίθμους, ακόμη και σε ανταγωνιστικές (*adversarial*) υλοποιήσεις του Κουτιού της Πανδώρας και του MSSC. Παρότι η ίδια *scenario-aware* προσέγγιση εφαρμόζεται τόσο για PA όσο και για NA πρότυπα, οι σταθεροί παράγοντες ενδέχεται να διαφέρουν λόγω της λιγότερο περιοριστικής φύσης των PA πολιτικών. Ωστόσο, τα αποτελέσματα δείχνουν ότι ακόμη και σε δύσκολα *bandit* περιβάλλοντα, είναι εφικτή η επίδοση κοντά στις γνωστές βέλτιστες προσεγγίσεις εκτός σύνδεσης, περιλαμβανομένης της βέλτιστης 4-προσεγγιστικής λύσης για το MSSC, επιβεβαιώνοντας τη δύναμη του συνδυασμού κυρτών χαλαρώσεων, *online learning* και ευέλικτων τεχνικών στοργγυλοποίησης.

4 Σύγκριση με Προηγούμενες Εργασίες

Η παρούσα εργασία βασίζεται σε θεμελιώδη αποτελέσματα της στοχαστικής βελτιστοποίησης, ειδικά όσον αφορά τα προβλήματα Κουτιού της Πανδώρας και MSSC. Προηγούμενες μελέτες εστίαστηκαν κυρίως στην στοχαστική περίπτωση, όπου οι τιμές και τα κόστη προέρχονται από γνω-

στές κατανομές, επιτρέποντας την ανάπτυξη αποδοτικών εκτός σύνδεσης αλγορίθμων και μερικώς προσαρμοστικών στρατηγικών [2, 8].

Αντιθέτως, εδώ εξετάζεται η *online* εκδοχή, όπου τιμές και κόστη ορίζονται ανταγωνιστικά σε πολλαπλούς γύρους. Το προτεινόμενο πλαίσιο αξιοποιεί χαλαρώσεις κυρτού τύπου (*convex relaxations*) και τεχνικές *online* κυρτής βελτιστοποίησης, προσφέροντας εγγυήσεις ακόμα και σε *bandit* περιβάλλον [1]. Παράλληλα, τα προτεινόμενα αποτελέσματα γενικεύουν τους αλγορίθμους του MSSC, επιτυγχάνοντας σταθερή προσεγγιστική απόδοση ακόμα και σε ματρωειδείς περιορισμούς — ένα σημείο στο οποίο παλαιότερες εργασίες είχαν περιορισμούς [3].

5 Εφαρμογές και Μελλοντική Έρευνα

Τα συμπεράσματα της μελέτης έχουν πρακτική σημασία σε διάφορους τομείς, όπως:

- **Διαχείριση πόρων:** Βελτιστοποίηση του κόστους ελέγχου στην παραγωγή και την ποιτική αξιολόγηση.
- **Αλγόριθμοι αναζήτησης:** Βελτίωση στρατηγικών για την εύρεση βέλτιστων λύσεων υπό αβεβαιότητα.
- **Online δημοπρασίες:** Προσαρμογή στρατηγικών πλειοδοσίας σε ανταγωνιστικά περιβάλλοντα.

Μελλοντικά, θα μπορούσε να διερευνηθεί η επέκταση του προτεινόμενου πλαισίου σε δυναμικά περιβάλλοντα, όπου οι περιορισμοί μεταβάλλονται με την πάροδο του χρόνου ή υπάρχουν πιο πλούσιοι μηχανισμοί ανατροφοδότησης στο *bandit* μοντέλο. Επιπλέον, μια άλλη ενδιαφέρουσα κατεύθυνση έρευνας είναι η βελτίωση των παραγόντων προσεγγιστικής επίδοσης σε πιο σύνθετα συνδυαστικά προβλήματα με πολλούς περιορισμούς.

6 Συμπεράσματα

Παρουσιάστηκε ένα ολοκληρωμένο πλαίσιο για τη διαδικτυακή (online) εκδοχή των προβλημάτων Κουτιού της Πανδώρας και *Min Sum Set Cover (MSSC)*, τα οποία αποτελούν δύο βασικούς πυλώνες στη στοχαστική βελτιστοποίηση. Με την αξιοποίηση κυρτών χαλαρώσεων και τεχνικών *online* κυρτής βελτιστοποίησης, προέκυψαν αποδοτικοί αλγόριθμοι που εγγυώνται σταθερό λόγο ανταγωνιστικότητας, τόσο σε συνθήκες πλήρους πληροφορίας όσο και σε σενάρια *bandit*. Οι αλγόριθμοι αυτοί επεκτείνουν τις δυνατότητες παλαιότερων προσεγγίσεων, ιδιαίτερα σε ανταγωνιστικά περιβάλλοντα και σε σύνθετους περιορισμούς όπως ματρωειδείς δομές [1].

A Παράρτημα

Σε αυτή την ενότητα μπορούν να συμπεριληφθούν πρόσθετες αποδείξεις, επεξηγήσεις ή συμπληρωματικά αποτελέσματα.

References

- [1] Evangelia Gergatsouli and Christos Tzamos. Online learning for min sum set cover and pandora’s box. In *International Conference on Machine Learning*, pages 7382–7403. PMLR, 2022.
- [2] Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.
- [3] Uriel Feige, László Lovász, and Prasad Tetali. Approximating min sum set cover. *Algorithmica*, 40:219–234, 2004.
- [4] Shai Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends® in Machine Learning*, 4(2):107–194, 2012.
- [5] Nikhil Bansal, Anupam Gupta, and Ravishankar Krishnaswamy. A constant factor approximation algorithm for generalized min-sum set cover. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1539–1545. SIAM, 2010.
- [6] Abraham D Flaxman, Adam Tauman Kalai, and H Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *arXiv preprint cs/0408007*, 2004.
- [7] Shai Shalev-Shwartz and Yoram Singer. A primal-dual perspective of online learning algorithms. *Machine Learning*, 69:115–142, 2007.
- [8] Shuchi Chawla, Evangelia Gergatsouli, Yifeng Teng, Christos Tzamos, and Ruimin Zhang. Pandora’s box with correlations: Learning and approximation. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1214–1225. IEEE, 2020.