

УДК 519.83

Д.Г. Павлов, М.В. Александрова, О.Р. Чертов

## ТЕОРЕТИКО-ІГРОВІ МОДЕЛІ ДЛЯ ВИЗНАЧЕННЯ МЕРЕЖЕВОГО ШАХРАЙСТВА В СИСТЕМІ КОНТЕКСТНОЇ РЕКЛАМИ

**Постановка проблеми.** Із розвитком Інтернету все частіше відбувається зіткнення декількох сторін, що мають різні або протилежні інтереси [1]: звичайні користувачі електронної пошти та такі, що навмисно розсилають рекламні повідомлення (спамери); легітимні користувачі та зломщики (хакери); пошукові мережі та власники сайтів, що використовуються в якості рекламних майданчиків. Оскільки математичною моделлю протистояння декількох сторін є гра, подібні ситуації прийнято розглядати в термінах теорії ігор.

Кількість учасників (гравців) та доходи окремих сторін в таких системах можуть бути дуже великими. Тому такі ігри виділяються в окремий клас – «Інтерактивні ігри з дуже великою кількістю гравців» (MMOG – massively multiplayer online games) [2]. Матеріальна зацікавленість сторін в інтерактивних іграх призводить до зростання рівня шахрайства в них.

Яскравим прикладом Інтерактивної гри з дуже великою кількістю гравців є система контекстної реклами. При обробці кожного пошукового запиту рекламна компанія в якості результатів пошуку надає посилання на найрелевантніші сайти та рекламні оголошення, які скоріше за все можуть зацікавити користувача. Обирання конкретних оголошень та їх ранжування по позиціях відбувається шляхом проведення узагальненого аукціону другої ціни (GSP – general second price auction) [3]. Учасниками такого аукціону є пошукова мережа, метою якої є максимізація власних прибутків, та рекламодавці, які прагнуть продемонструвати свій сайт найбільшій кількості цільової аудиторії. При розміщенні оголошень на сайтах з'являється ще один учасник гри – власник сайту. Оскільки він отримує деяку платню за кожен перехід по оголошенню, розміщеному на його сайті, природним прагненням для нього є збільшення кількості кліків (натискань по рекламних повідомленнях). Головною особливістю аукціонів Інтернет-реклами є висока динамічність. Кожен рекламодавець може в будь-який час змінити свою ставку за клік, внаслідок чого гра не може бути стійкою.

Розглянута система є вразливою до такого різновиду мережевого шахрайства як склікування, коли відбувається штучна генерація переходів по оголошенню з метою розтрати рекламного бюджету рекламодавця (конкуренція між рекламодавцями за вищі позиції при показі) або збільшення власних прибутків (властиво власникам сайтів). Відомо, що рівень недійсних кліків в мережі є достатньо високим, наприклад, по даним компанії Adometry в 4-му кварталі 2010 р. він склав 19,1 % від усіх переходів [4]. Тому розробка нових методів боротьби із склікуванням є актуальною задачею.

**Аналіз публікацій по темі дослідження.** Багато дослідників працюють над задачею формалізації відносин учасників процесу контекстної реклами в термінах теорії ігор. Приміром, в роботі [3] була описана та досліджена модель узагальненого аукціону другої ціни, що наразі в чистому вигляді використовується в системі Yahoo! та із невеликими змінами в пошуковій мережі Google. В статті [5] розглянуто динаміку досягнення рівноважного стану декількома рекламодавцями.

Окреме місце займають дослідження, націлені на розробку таких моделей аукціону, що є найбільш вигідними для пошукових мереж. В роботі [6] був запропонований алгоритм визначення позицій оголошень для системи контекстної реклами AdWords пошукової мережі Google. В [7] було показано, що використання елементів випадкового вибору замість строгого розташування оголошень в порядку

зменшення їх ставки за клік може позитивно відобразитись на доходах пошукових мереж.

Автори роботи [8] розглядають можливу поведінку нового учасника процесу Інтернет-реклами – постачальника Інтернет-послуг. Останні володіють специфічною інформацією щодо уподобань своїх користувачів та можуть або формувати коаліцію із пошуковими мережами задля досягнення найкращих результатів, або інтерактивно змінювати результати пошуку (зокрема, оголошення). Тому постачальники Інтернет-послуг стають активною частиною системи рекламування. Обирання конкретної стратегії поведінки залежить від того, яка з них принесе більші прибутки.

Статті [9, 10] присвячені моделюванню можливих стратегій поведінки власників сайтів. Автори визначають умови, за яких власникам сайтів вигідно проводити легітимні дії або генерувати штучні кліки. В роботі [10] також будується аналогічна модель для процесу спілкування оголошень конкурентів.

Головною особливістю всіх зазначених вище робіт є те, що розроблені теоретико-ігрові моделі використовуються для аналізу ситуації, що склалася в системі контекстної реклами, з метою визначення шляхів модифікації аукціону або стратегічних дій пошукових мереж таким чином, щоб знизити зацікавленість можливих шахраїв в процесі спілкування. Для цього необхідно зробити затрати на генерацію недійсних кліків вищими за можливий від них прибуток.

Проте, зрозуміло, що для ситуацій масштабного спілкування прибутки від шахрайських дій можуть бути достатньо великими. Цей факт забезпечує наявність стимулу до подальшого шахрайства в мережі.

На відміну від зазначених вище робіт автори статті вважають, що теоретико-ігрові моделі також можуть використовуватись для визначення наявності спілкування в мережі.

**Метою статті** є аналіз існуючих теоретико-ігрових моделей системи контекстної реклами на предмет можливості їх подальшого використання для боротьби із мережевим шахрайством.

**Постановка задачі. Узагальнений аукціон другої ціни.** Оскільки результати проведення аукціону визначають позиції рекламних оголошень, що є одним з найголовніших елементів в системі Інтернет-реклами, розуміння цього процесу є необхідним для побудови подальших моделей. Розглянемо головні аспекти базової моделі аукціону другої ціни, що наразі використовується системою Yahoo! [3].

Введемо наступні позначення. Нехай  $N$  – кількість позицій на екрані;  $K$  – кількість рекламодавців, що претендують на показ свого оголошення за деяким ключовим словом;  $\alpha_i$  – математичне очікування кількості кліків по оголошенню, розташованому на позиції  $i$ ;  $s_k$  – корисність одного переходу для рекламодавця  $k$ . Тоді вираз  $\alpha_i s_k$  представлятиме собою дохід, який рекламодавець  $k$  отримає в результаті перебування його оголошення на позиції  $i$ . Якщо позначити через  $p_i$  плату рекламодавця за перебування на позиції  $i$ , тоді його чистий прибуток становитиме

$$u_i^k = \alpha_i s_k - p_i. \quad (1)$$

Позначимо через  $b_i$  максимальну суму, яку рекламодавець згоден платити за один перехід по його оголошенню. Величина  $b_i$  є ставкою в рамках даного аукціону і може бути змінена в будь-який момент. Пошукова мережа розміщує оголошення в порядку зменшення значення  $b_i$ . Якщо декілька рекламодавців приймають однакову величину ставки, їх оголошення можуть розміщуватись випадково. Однак, на практиці використовується таке правило: позиція оголошення буде меншою для того

рекламодавця, який першим прийняв ставку. Цей підхід є аналогом правила проведення аукціонів: хто першим здається, той програє.

Якщо на оголошення було натиснуто, то з рекламодавця знімається плата, що дорівнює ставці наступного за ним гравця. Іншими словами, загальна плата пошуковій системі становить  $p_i = \alpha_i b_{i+1} \quad \forall i \in \{1, \dots, \min(N, K)\}$ , тобто, використовуючи рівняння (1), можна записати, що  $u_i = \alpha_i (s_i - b_{i+1})$ .

Оскільки кількість гравців в подібній системі може бути дуже великою, встановлення рівноважного стану в грі є достатньо важкою задачею. Тому розглядається локально-стійка рівновага, тобто такий стан, коли рекламодавець не може збільшити свого виграшу за рахунок зміщення гравця, розташованого на 1 позицію вище. Тобто умовою рівноваги є нерівність  $u_i^{g(i)} \geq u_{i-1}^{g(i)}$ , де  $g(i)$  – ідентифікатор гравця на позиції  $i$ .

У випадку врахування якості рекламного оголошення кількість кліків, яку отримує рекламодавець, перебуваючи на позиції  $i$ , становитиме  $\alpha_i \beta_k$ , де  $\beta_k$  – показник якості. Необхідною та достатньою умовою локально-стійкої рівноваги в цьому випадку буде нерівність

$$\alpha_i \beta_{g(i)} (s_{g(i)} - b_{i+1}) \geq \alpha_j \beta_{g(i)} (s_{g(i)} - b_{j+1}),$$

де  $i > j$ .

Тобто  $\alpha_i (s_{g(i)} - b_{i+1}) \geq \alpha_j (s_{g(i)} - b_{j+1})$ , а, отже, рівновага не залежить від  $\beta_k$ .

**Аукціон Google.** В системі Google оголошення розміщуються в порядку зменшення значення рангу, що становить  $b_i \gamma_{g(i)}$ , де  $\gamma_k$  – показник якості рекламного оголошення. Плата за клік  $x_i$  становить мінімально необхідну суму для того, щоб виконувалась умова

$$\gamma_{g(i)} x_i \geq \gamma_{g(i+1)} b_{g(i+1)},$$

тобто  $x_i \geq \frac{\gamma_{g(i+1)} b_{g(i+1)}}{\gamma_{g(i)}}$ .

Необхідною та достатньою умовою рівноваги є нерівність

$$\forall i, j > i, \quad \alpha_i \beta_{g(i)} \left( s_{g(i)} - \frac{\gamma_{g(i+1)} b_{g(i+1)}}{\gamma_{g(i)}} \right) \geq \alpha_j \beta_{g(i)} \left( s_{g(i)} - \frac{\gamma_{g(j+1)} b_{g(j+1)}}{\gamma_{g(i)}} \right),$$

або  $\alpha_i (\gamma_{g(i)} s_{g(i)} - \gamma_{g(i+1)} b_{g(i+1)}) \geq \alpha_j (\gamma_{g(i)} s_{g(i)} - \gamma_{g(j+1)} b_{g(j+1)})$ .

Тобто умова рівноваги для системи Google співпадає із умовою рівноваги узагальненого аукціону другої ціни, де встановлені ставки  $\{\gamma_k b_k\}$ , а вартості кліків для рекламодавців дорівнюють  $\{\gamma_k s_k\}$ .

**Стратегії поведінки власників сайтів.** Оскільки кожен клік по оголошенню, розміщеному на сайті, що є рекламним майданчиком, приносить його власникові певний прибуток, то загальне збільшення кількості переходів по оголошенню є природною стратегією такого гравця. Збільшити кількість переходів можна двома способами: легітимним та шахрайським. Перший спосіб передбачає покращення якості сайту, розроблення цікавого контенту, залучення нових відвідувачів та утримання «старої» аудиторії. Подібні дії передбачають відповідні витрати. Другий спосіб – генерування штучних переходів.

Оскільки пошукові мережі розроблюють спеціальні системи захисту від склікування, процес генерації недійсних кліків також потребує певних витрат. До того ж, із збільшенням кількості недійсних переходів, підвищується ймовірність того, що шахрайські дії будуть викриті, а, отже, знижується їх вартість для рекламодавця.

Умова легітимної поведінки власника сайту може бути записана в формі нерівності

$$Px - \beta x \geq Pr - \alpha r - V(r),$$

де  $P$  – дохід власника сайту від одного кліку,  $x$  – кількість дійсних кліків, які можуть бути здійснені на сайті,  $\beta$  – витрати власника сайту, що припадають на один дійсний перехід,  $r$  – кількість недійсних кліків,  $\alpha$  – витрати на генерацію одного недійсного кліку,  $V(r)$  – зростаюча функція, що визначає можливі втрати власника сайту у випадку викриття шахрайства.

**Висновки та перспективи подальшого дослідження.** В даній роботі розглянуті існуючі теоретико-ігрові моделі системи контекстної реклами. Під час рекламування в мережі Інтернет, зазвичай, взаємодіють наступні сторони: пошукова мережа, рекламодавці та власники сайтів. Кожна із сторін прагне максимізувати свої прибутки, що і визначає стратегію їх поведінки.

Перспективами подальшого дослідження є розробка теоретико-ігрових моделей поведінки кожного із гравців в умовах використання ними легітимної та шахрайської стратегій. Аналізуючи реальну поведінку кожного суб'єкта із використанням розроблених моделей можна буде в подальшому зробити висновки відносно наявності склікування в системі.

Особливий інтерес представляє введення в систему нового гравця олігополіста-альтруїста [11], метою якого є не власне збагачення, а задоволення деяких потреб соціально-незахищених груп суспільства. Такі організації на сучасному етапі займають вагомий позицію в економіках деяких країн Європи, Канади та Японії.

#### ЛІТЕРАТУРА:

1. Kanatarciogly M. A Game Theoretic Approach to Adversarial Learning / M. Kanatarciogly, B. Xi, C. Clifton. – Purdue University, Department of Statistics, Technical Report /05-06, 2005. – 12 p.
2. Virtual Worlds and Fraud: Approaching Cybersecurity in Massively Multiplayer Online Games / J. Bardzell, M. Jakobsson, S. Bardzell et al. // Proceedings of DiGRA 2007 Conference. – 2007. – P. 451–742.
3. Edelman B. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords / B. Edelman, M. Ostrovsky, M. Schwarz // American Economic Review. – 2007. – Vol. 97. – P. 242–259.
4. Click Fraud Rate Drops to 19.1 Percent in Q4 2010 [Електронний ресурс]. – Режим доступу: <http://www.adometry.com/media/press/release.php?id=1>.
5. Bu T. Dynamics of strategic manipulation in ad-words auction / T. Bu, X. Deng, Q. Qi // Proc. of the 3rd Workshop on Ad Auctions, Banff, Canada. – 2007.
6. AdWords and Generalized On-line Matching / A. Mehta, A. Saberi, U. Vazirani et al. // 46th Annual IEEE Symposium on Foundations of Computer Science. – 2005. – P. 264–273.
7. Maille P. On the interest of introducing randomness in Ad-word auctions / P. Maille, B. Tuffin // Information and Communication Technology. – 2010. – Vol. 327. – P. 229–240.

8. Security Games in Online Advertising: Can Ads Help Secure the Web? / N. Vratonjic, M. Raya, J.-P. Hubaux et al. // Workshop on the Economics of Information Security (WEIS 2010). – 2010.
9. Asdemir K. An Economic Model of Click Fraud in Publisher Networks / K. Asdemir, O. Yurtseven, M. A. Yahya // International Journal of Electronic Commerce. – 2009. – Vol. 13. – No. 2. – P. 61–89.
10. Wilbur K. Click Fraud / K. Wilbur, Y. Zhu // Marketing Science. – 2009. – Vol. 28. – No. 2. – P. 293–308.
11. Mixed oligopoly with consistent conjectures / V. Kalashnikov, V. Bulavsky, I. Kalashnykova et al. // European Journal of Operational Research. – 2010. – Vol. 210. – No. 3. – P. 729–735.

ПАВЛОВ Дмитро Геннадійович – аспірант кафедри прикладної математики Національного технічного університету України «Київський політехнічний інститут».

Наукові інтереси:

– моделювання складних систем, інтелектуальний аналіз даних, Інтернет-реклама.

АЛЕКСАНДРОВА Маргарита Володимирівна – магістрантка кафедри прикладної математики Національного технічного університету України «Київський політехнічний інститут».

Наукові інтереси:

– інтелектуальний аналіз даних, вейвлет-перетворення.

ЧЕРТОВ Олег Романович – к.т.н., доцент кафедри прикладної математики Національного технічного університету України «Київський політехнічний інститут»

Наукові інтереси:

– інформаційні технології, інтелектуальний аналіз даних, метадані;  
– вейвлет-перетворення, знеособленість та анонімність даних.