

### *Література*

1. Тогобицкая Д.Н. Формирование интегрированной базы экспериментальных данных о составе и технологических параметрах технологии десульфурации зернистым магнием /Д.Н.Тогобицкая, Л.С. Молчанов, А.С. Вергун, В.Г. Кисляков, Ю.М. Лихачев // *Литье. Металлургия. 2019: Материалы XIV МНПК, Запорожье, 21-23 мая 2019.* – Запорожье: «ЗТПП». – 2019. – С.360-362.

2. Тогобицкая Д.Н. Подсистема «ВОЧ» в решении задач выбора рациональной технологии получения качественной металлопродукции в сложившихся сырьевых и технологических условиях /Д.Н.Тогобицкая, А.С.Вергун, Л.С.Молчанов, В.Г.Кисляков, Ю.М.Лихачев, Н.Е.Ходотова //Сб.н.тр. ИЧМ «Фундаментальные и прикладные проблемы черной металлургии». – Днепр. -2019. – Вып.33.

## **Дослідження методів моделювання і інформаційних технологій для реалізації збільшення відвідувань сайту**

**Твердохліб І. А., Гурковська С. С.**

*Донбаська державна машинобудівна академія*

Важливим завдання кожного сайту з відео-контентом по типу серіалів є залучення активних користувачів, які усвідомлено обрали б цей сайт для перегляду необхідного їм контенту. Зазвичай користувач може обрати сайт не зважаючи на якість відео чи звуку. Для покращення якості контенту, який споживають користувачі виникає потреба ретельного аналізу наповнення сайту. Також з'являється необхідність аналізу дій користувачів, з метою визначення необхідних змін в функціоналі сайту, або якості контенту. Такі дані використовуються з метою редизайну сайту. Для таких робіт долучають багато різних спеціалістів. При цьому ціна таких послуг може сягати десятки тисяч гривень.

З метою зменшення витрат на проведення робіт з аналітики була розроблена математична модель поведінки системи «відвідувач-сайт». За основу приймалася така поведінка. У кожен момент часу система може перебувати в одному з  $K + 1$  станів  $S$  з номерами  $k = 0, 1, \dots, K$ , де  $S_0$ - перебування відвідувача на будь-якому іншому сайті Інтернету,  $S_1$ - перебування користувача на головній (першій) сторінці даного сайту,  $k$  – номер сторінки сайту, що розглядається. Початковий розподіл ймовірностей при цьому має вигляд [1]:

$$P_0(0) = 1, P_i(0) = 0, i = 1, 2, \dots, K$$

Оскільки користувач може потрапити в будь-який стан  $S_0 \dots S_k$  тільки з іншого сайту або набравши адресу однієї зі сторінок  $1 \dots K$  в адресному рядку браузера. При запропонованому використанні напівмарківської моделі з метою оптимізації структури та навігації або побудови оптимальної рекламної стратегії достатньо задати матрицю однокрокових перехідних ймовірностей  $(\pi_{ij})$ ,  $i = 0 \dots K$ ,  $j = 0 \dots K$  та безліч середніх часів перебування в станах  $\langle T_i \rangle$   $i = 0 \dots K$ . Для цього було необхідно організувати спостереження за візитами на сайт протягом періоду, достатнього для отримання статистично стійкого ансамблю результатів. Тривалість цього періоду залежала від інтенсивності відвідувань.

У ході спостереження реєструвалися:

- а) переходи з одних сторінок на інші (події, яким відповідають елементи матриць  $(\pi_{ij})$  з викресленими першими стовпцем та рядком);
- б) приходи на сайт (події, яким відповідають елементи рядка  $(\pi_{0j})$ );
- в) відходи з сайту (події, яким відповідають елементи стовпця  $(\pi_{0j})$ );
- г) часи перегляду сторінок, які відповідають часам перебування  $T_i$ ,  $i \in [1, K]$  станах  $S_1 \dots S_k$  [2].

При цьому як  $\langle T_0 \rangle$  розглядался середній час сесії користування Інтернетом за вирахуванням середнього часу, проведеного на сайті, що розглядався. Після отримання статистичних оцінок параметрів моделі  $((\pi_{0j})_{ij})$  та  $\langle T_i \rangle$  здійснювалося рішення системи (1) щодо  $P_i$ :

$$P_k = \sum_{j=0}^K \pi_{jk} P_j, k = 0, 1, \dots, K \quad (1)$$

і обчислення ймовірностей  $Q_i$  за формулою (2), що являють собою фінальні ймовірності перебування у станах, зважені з відносними середніми безумовними часами перебування у станах.

$$Q_i = P_i \langle T_i \rangle \sum_{j=0}^K P_j \langle T_j \rangle \delta^{-1} \quad (2)$$

$Q_i$  - змістовні параметри, що дозволяють ефективно вирішувати завдання ранжирування сторінок сайту для оптимального розміщення реклами, а також завдання загальної порівняльної оцінки відвідуваності сторінок сайту для користувачів, яка у багатьох випадках синонімічна корисність.

Важливими характеристиками веб-сторінок є також величина часу та число переходів, необхідних для досягнення відвідувачем конкретної сторінки з головної сторінки. Знання цих характеристик дозволяє об'єктивно оцінювати доступність сторінок для відвідувача та цілеспрямовано оптимізувати структуру та навігацію сайту. Визначення найбільш ймовірний час та кількість переходів, необхідних для досягнення деякої сторінки  $j$  з головної сторінки виконувалося за допомогою розподілу ймовірностей:

$$g_{ij} = pn(t) = n, S(t) = S_j * S_i, n(0) = 0 \quad (3)$$

де  $g_{ij}(n, t)$  - спільна ймовірність того, що потрібно  $n$  переходів і час  $t$  для досягнення вперше стану  $S_j$  якщо при  $t = 0$  система перебувала у стані  $S_1$ . Ймовірності  $Q$  та  $g$  пов'язані співвідношенням (3):

$$Q_{ij}(n + 1, t) = \sum_{m=0}^n \int_0^t g_{ij}(m + 1, \tau) Q_{ij}(n - m, \tau) d\tau \quad (4)$$

де  $Q_{ij}(n, t)$  - спільна ймовірність того, що система знаходиться в стані  $S_j$  в момент  $t$  і що вона зробила  $n$  переходів за умови, що в момент часу  $t = 0$  вона перебувала в стані  $S_i$ .

Вирішуючи (4) одержали шуканий розподіл  $g_{1j}(n, t)$ . Наприклад, застосування перетворення Лапласа до (4) призводить до системи  $n$  рівнянь алгебри:

$$Q_{1j}(n + 1, s) = \sum_{m=0}^n g_{ij}(m + 1, s) Q_{ij}(n - m, s) \quad (5)$$

Рішення системи (4.1) дозволяє визначити  $g^*_{lj}(n,s)$  та після зворотного перетворення Лапласу  $g_{lj}(n,t)$ . Це рішення вимагає наявності статистичних оцінок  $Q^*_{lj}(n,t)$  для всіх можливих поєднань  $n$  і  $t$ , тому щоб зробити процедуру практично здійсненої континуум значень  $t$  замінюється послідовністю відліків  $t_i$ , взятих з кроком, що забезпечує прийнятну практичну точність відтворення розподілів  $g_{lj}(n,t)$ . Найбільш ймовірні час і кількість переходів, необхідних для досягнення певної сторінки  $j$  з головної сторінки, визначаються як моди  $g_{lj}(n,t)$  за аргументами  $t$  і  $n$  відповідно. Можна реалізувати [3], ґрунтуючись на розподілах інтервально-перехідних ймовірностей, які визначаються як розв'язання системи лінійних інтегральних рівнянь:

$$Q_{ij}(t) = \delta_{ij}H_i(t) + \sum_{k=0}^K \pi_{ik} \int_0^t f_{ik}(\tau) Q_{kj}(1 - \tau) d\tau \quad (5)$$

де  $Q_{ij}(t)$  - умовна ймовірність того, що в момент часу  $t$  система знаходиться в стані  $S_j$ , якщо в момент часу  $t = 0$  вона знаходилася в стані  $S_i$ ,  $\delta_{ij}$  символ Кронекера,  $\delta_{ij} = 1$ , якщо  $i = j$  та  $\delta_{ij} = 0$ , якщо  $i \neq j$ ;  $H_i(t)$  - ймовірність того, що система не залишить стану  $S_i$  до моменту  $t$ . Знаючи  $Q_{ij}(t)$  для стану  $S_i$  можна організувати послідовність показів різних об'єктів «відстроченої» реклами, кожен із яких містить.

Таким чином була отримана математична модель поведінки системи «відвідувач-сайт». Вона дозволяє моделювати усі події користувача на сайті, як, наприклад, його переходи по різних сторінкам. Це значно скоротить час та матеріальні ресурси для виконання аналізу, який потрібен для редизайну сайту (з метою збільшення його відвідування).

#### Література

1. А.Л.Горбунов «Марківські моделі відвідуваності веб-сайтів»
2. Зі списку статті посилання №1 Айвазян С.А., Мхітарян В.С. Прикладна статистика та основи економетрики-М: ЮНІТІ,1998.
3. Adam Clarke. SEO 2021 Learn Search Engine Optimization With Smart Internet Marketing Strategies: Learn SEO with Smart Internet Marketing Strategies / Amazon Digital Services LLC - KDP Print US, 2020 – 259 p.