

גירסה 1.00 – 8.6.2005



# אלגוריתם TrustRank

ניר אדר

מסמך זה הורד מהאתר <http://www.underwar.co.il>.  
אין להפיץ מסמך זה במדיה כלשהי, ללא אישור מפורש מאת המחבר.  
מחבר המסמך איננו אחראי לכל נזק, ישיר או עקיף, שיגרם עקב השימוש במידע המופיע  
במסמך, וכן לנכונות התוכן של הנושאים המופיעים במסמך. עם זאת, המחבר עשה את מירב  
המאמצים כדי לספק את המידע המדויק והמלא ביותר.

כל הזכויות שמורות לניר אדר

Nir Adar

Email: [underwar@hotmail.com](mailto:underwar@hotmail.com)

Home Page: <http://www.underwar.co.il>

אנא שלחו תיקונים והערות אל המחבר.

## תוכן עינינים

3.....	GOOGLE TRUSTRANK
3.....	1. מבוא – הצגת הבעיה
4.....	2. הרעיון מאחורי האלגוריתם
5.....	3. הכנה להסבר על האלגוריתם
5.....	3.1 מידול רשת האינטרנט
5.....	3.2 אלגוריתם PageRank
7.....	4. אמון
7.....	4.1 פונקציית ההערכה $O$
8.....	4.2 חישוב האמון
9.....	4.3 החלשות האמון ודעיכת האמון
10.....	5. אלגוריתם TRUSTRANK
11.....	6. בחירת קבוצת השורש
11.....	6.1 בחירת קבוצת שורש אקראית
11.....	6.2 בחירת קבוצת שורש לפי PageRank
12.....	6.3 בחירת קבוצת השורש לפי PageRank הפוך
13.....	7. סיכום ומסקנות

## Google TrustRank

### 1. מבוא – הצגת הבעיה

דפי ספאם רבים ברשת האינטרנט משתמשים בטכניקות שונות על מנת לקבל ניקוד גבוה יותר מהמגיע להם בתוצאות מנועי החיפוש.

ניתן לסווג את דפי הספאם לשתי קבוצות עקריות:

1. **דפים המכילים טקסט וקישורים בלתי נראים בתוכם.** על ידי הוספת הטקסט והקישורים הבלתי נראים, האתר נחשב רלוונטי עבור המילים הנוספות הבלתי נראות, למרות שבפועל אינו קשור אליהן.
2. **קבוצת אתרים גדולה שנוצרת במיוחד על מנת להצביע אל דף מסויים.** מכיוון שכל מנועי החיפוש כיום מחשיבים קישור כאלמנט של הצבעה עבור הדף אליו מקשרים, ערכו של הדף ומיקומו במנועי החיפוש עולה באופן לא טבעי.

בני אדם יכולים לזהות דפים אלו, אך כמות הדפים באינטרנט גדולה מכדי שבני אדם יוכלו לעבור עליה ולסנן את תוצאות הספאם עבור מנועי החיפוש.

המטרה של מנועי החיפוש, מאז ומתמיד, הינה לבנות אלגוריתמים אוטומטיים שיאפשרו להם להבדיל בין דפים טובים לדפי ספאם, ולקדם לראש התוצאות את הדפים הטובים, הלגיטימיים ביותר לנושא אותו מחפש הגולש. **TrustRank** הינו אלגוריתם חדש (2004) המנסה להשיג תוצאות טובות יותר בסינון ספאם מהאלגוריתמים הקודמים (דגש במיוחד על אלגוריתם PageRank).

השם *TrustRank* נרשם על ידי גוגל כסימן רשום במרץ 2005, וכעת הרבה מקדמי אתרים רואים בכך סימן שהם הולכים לאמץ גרסה כלשהי של האלגוריתם בקרוב, אם הם לא משתמשים כבר בגרסה כזו.

מסמך זה בא להציג לקורא את האלגוריתם החדש. המסמך מתבסס בעיקר על המאמר משנת 2004 – "Combating Web Spam with TrustRank".

## 2. הרעיון מאחורי האלגוריתם

זיהוי ספאם באמצעות כלים אוטומטיים הינו בעייתי. אנשים לומדים כיצד הכלים המזהים פועלים, ומוצאים את הדרך לגרום לאתרים שלהם להראות חוקיים עבור הכלים האוטומטיים.

TrustRank משלב מעט עבודה של צוות אנושי, על מנת להפוך את הזיהוי האוטומטי ליעיל יותר. האלגוריתם מגדיר קבוצה ראשונית של דפים, **קבוצת השורש**, אשר עבורם יש צורך לקבוע באופן ידני האם הם **דפים טובים** או **דפים רעים** (ספאם). צוות מומחים אנושי בודק את קבוצת דפים זו ומזין את התוצאות אל האלגוריתם (טוב/רע). על פי קבוצה ראשונית זאת, ומבנה הקישורים בין הדפים באינטרנט, האלגוריתם מחליט עבור כל הדפים האחרים ברשת האינטרנט האם הם דפים טובים או רעים.

TrustRank מסתמך על ההנחות הבאות:

- דפים טובים מקשרים לעתים רחוקות בלבד אל דפים רעים. לעומתם דפים רעים מקשרים לרוב אל דפים טובים, על מנת לגרום למנועי החיפוש לחשוב שהם אתרים לגיטימיים.
- קיים יחס הפוך בין מספר הקישורים בעמוד מסויים לבין החשיבות שיוצר העמוד נותן לכל קישור.
- ככל שמתרחקים מהאתרים הטובים שנבדקו ידנית, כך גדל הסיכוי שהאתר שאליו נגיע הינו אתר רע. לכן, בדומה ל-PageRank קיים מנגנון דעיכה בהעברת האמון בין הדפים.

על מנת לבדוק את יעילות השיטה, כותבי המאמר המקורי בחרו קבוצה של כ-200 אתרים כקבוצת השורש. הם הצליחו להראות שעל ידי 200 אתרים בלבד שנבדקו באופן ידני, הם הצליחו לסנן חלק משמעותי מאתרי הספאם ברשת האינטרנט.

TrustRank מזכיר את אלגוריתם PageRank הוותיק של גוגל, אך עם שינוי אחד עקרוני. אלגוריתם PageRank מכריז על שוויון בין כולם ברשת האינטרנט – אין קישור שבאופן עקרוני שווה הרבה יותר מקישור אחר. TrustRank, לעומת זאת, מציג גישה שונה. האתרים הנבחרים כקבוצת השורש הינם מקור הכוח של האינטרנט. הם אלו שעל ידי הצבעה שלהם, קובעים למעשה מי "קרוב לצלחת" וראוי לאמון.

### 3. הכנה להסבר על האלגוריתם

#### 3.1 מידול רשת האינטרנט

על מנת לדון באלגוריתם, צריך מודל לרשת האינטרנט. כמו במקרים רבים, המודל שנבחר הינו גרף מכוון  $G(V, E)$ , בו כל דף הינו צומת, וקישור בין דף  $p \in V$  לדף  $q \in V$  הינו קשת  $(p, q) \in E$ . לצומת יכולים להכנס מספר קשתות (קישורים) וכן יכולים לצאת ממנו מספר כלשהו של קישורים. מספר הצמתים (הדפים) הכולל יסומן ב-  $N$ .

עבור דף  $p$  אנחנו יכולים לדבר על **מספר הקישורים הנכנסים אליו**, שיסומן ב-  $i(p)$ , וגם על **מספר הקישורים היוצאים ממנו**, שיסומן  $w(p)$ .

נגדיר **מטריצת מעברים** שתייצג את הדפים ברשת האינטרנט. המטריצה תיבנה על ידי המשוואה הבאה:

$$T(p, q) = \begin{cases} 0 & \text{if } (q, p) \notin E \\ 1/w(q) & \text{if } (q, p) \in E \end{cases}$$

#### 3.2 אלגוריתם PageRank

PageRank היה האלגוריתם המקורי בו גוגל השתמש על מנת לדרג אתרים, ובו נעשה שימוש (מופחת אומנם) עד היום. האלגוריתם מתבסס על ההנחה שדף אליו יש קישורים רבים – הוא ככל הנראה חשוב יותר. כל קישור אל האתר מהווה "הצבעת אמון". חישוב ה-PageRank עבור אתר נעשה בצורה הבאה:

$$PR(p) = \alpha \cdot \sum_{q:(q,p) \in E} \frac{PR(q)}{w(q)} + (1-\alpha) \cdot \frac{1}{N}$$

כאשר  $\alpha$  זהו מקדם דעיכה שערכו לפי האלגוריתם המקורי הוא 0.85.

על ידי שימוש ב- $T$ , אנחנו מסוגלים לייצג את ה-PageRank בכל רשת האינטרנט כווקטור:

$$PR = \alpha \cdot T \cdot PR + (1 - \alpha) \cdot \frac{1}{N} \cdot 1_N$$

כאשר  $1_N$  היא מטריצת היחידה ו- $PR$  הוא ווקטור כל האתרים. נשים לי כי תוצאת החישוב הינה ווקטור.

מטריצת היחידה בעצם נותנת משקל זהה לכל אחד מהדפים באינטרנט, באופן בסיסי, כאשר הדפים שמקבלים בסופו של דבר את ה-PageRank הגבוה ביותר הם אלו שהכי הרבה קישורים מתועלים אליהם.

ניתן ליצור במקום מטריצת היחידה מטריצה אחרת, **מטריצה משוחדת**, שתיתן ערך התחלתי שונה לכל דף. ה-PageRank המשוחד יחושב על ידי הנוסחה הבאה:

$$r = \alpha \cdot T \cdot r + (1 - \alpha) \cdot d$$

כאשר  $d$  זהו ווקטור בעל  $N$  איברים אשר כל איבריו אי שליליים, וסכומם הכולל הינו 1.

פעולה זו הינה בדיוק פעולתו של TrustRank. TrustRank מגדיר, בדרך שנראה בהמשך, מטריצה חדשה בה נשתמש לצורך חישוב דירוג הדף.

## 4. אמון

4.1 פונקצית ההערכה  $O$ 

כפי שצוין, כל דף הינו דף טוב (דף לגיטימי) או דף רע - ספאם שנוצר על מנת להטעות מנועי חיפוש. נגדיר את הפונקציה  $O$  המקבלת דף  $p \in V$  ומחזירה ערך בוליאני – האם הדף טוב או רע. ההנחה שכל דף עבורו מחושבת הפונקציה  $O$  נבדק על ידי בן אדם שקובע באופן מוחלט האם מדובר בדף לגיטימי או לא.

$$O(p) = \begin{cases} 0 & \text{if } p \text{ is bad} \\ 1 & \text{if } p \text{ is good} \end{cases}$$

כאמור – לא נרצה לחשב את  $O$  עבור כל הדפים באינטרנט מכיוון שאין מספיק כ"א ותקציב כדי לעשות זאת. המטרה היא לחשב את  $O$  עבור קבוצה נבחרת של דפים, שיהוו את קבוצת השורש שלנו.

לפיכך נשתמש בהנחת הבידוד של הקבוצה הטובה: דפים טובים מקשרים לעתים רחוקות בלבד אל דפים רעים. ההגיון הוא כזה: דפים רעים הם דפים שכל המטרה שלהם היא לעבוד על מנועי חיפוש ואין להם תוכן אמיתי, ולכן לדפים הטובים אין אינטרס לקשר אליהם. לעתים דפים לגיטימיים עשויים לקשר לאתרים רעים (הערת ספאם בתוך בלוג לגיטימי, למשל), אולם לרוב ההנחה מתקיימת.

נשים לב כי הנחת הבידוד לא עובדת לכיוון השני. דפי ספאם עלולים לקשר לדפים טובים, על מנת לנסות לשכנע את מנועי החיפוש כי הם אכן לגיטימיים.

כדי להחליט האם דף  $p$  שאינו בקבוצת השורש  $S$  הוא טוב או רע, נגדיר פונקצית הערכה

$T$  המחזירה ערך בתחום  $[0..1]$  המציין מה הסיכוי ש- $p$  הוא דף רע (0) או טוב (1):

$$T(p) = \Pr(O(p) = 1)$$

נשים לב שהפונקציה אינה אומרת אם הדף הוא טוב או רע, אלא מה הסיכוי שהוא יהיה דף טוב. למרות שאין לנו קביעה מוחלטת האם הדף טוב או רע, עדיין שימושי להשתמש בפונקציה הזו.

## 4.2 חישוב האמון

נציג כעת את הפונקציה  $T$  שתשמש להערכה. נתחיל בהצגת גרסה פשוטה של הפונקציה לצורך ההבנה, ואז נציג את הגישה בה נשתמש בפועל.

נבחר קבוצה אקראית  $S$  של  $L$  דפים עליהם נחשב את  $O$ . קבוצה זו היא קבוצת השורש שלנו. נסמן את הצמתים הטובים בקבוצה ב- $S^+$  ואת הרעים ב- $S^-$ .

הפונקציה הראשונה שנציע כ- $T$  הינה הפונקציה הבאה:

$$T_0(p) = \begin{cases} O(p) & \text{if } p \in S \\ 1/2 & \text{otherwise} \end{cases}$$

עבור כל הדפים בקבוצת השורש ערכם ידוע, וערך האמון של כל הדפים האחרים הוא  $1/2$  – לא לכאן ולא לכך.

פונקציה זו מספקת את הכיוון ההתחלתי, אולם היא איננה עונה על הצרכים שלנו. היא איננה משתמשת בהנחה שהצגנו כי דפים טובים מקשרים רק אל דפים טובים אחרים. לפיכך, נציג שיפור לפונקציה: הפונקציה תחליט שצמתים הנמצאים במרחק של עד  $M$  (עבור  $M$  קבוע) קשתות מצומת טוב, הינם גם כן טובים:

$$T_M(p) = \begin{cases} O(p) & \text{if } p \in S \\ 1 & \text{if } p \notin S \text{ and } \exists q \in S^+ : q \xrightarrow{M} p \\ 1/2 & \text{otherwise} \end{cases}$$

כאשר  $q \xrightarrow{M} p$  משמעו שקיים מסלול באורך של עד  $M$  בין  $q$  ל- $p$ , שאינו מכילים צמתים רעים מקבוצת השורש.

חשיבותה של פונקציה זו: הפונקציה מאפשרת לנו להגדיל בצורה משמעותית ביותר את קבוצת האתרים אותם אנחנו מחשיבים כ-"טובים".

על ידי הפונקציה  $T_M(p)$  נוכל לבנות ווקטור המכיל את הערכים עבור כל דפי האינטרנט. ווקטור זה ישמש אותנו כ- $d$  עבור חישוב ה-PageRank המשוחד שהצגנו.



### 4.3. החלשות האמון ודעיכת האמון

לאחר הגדרת האתרים הטובים כפי שראינו בפרק הקודם, מתחיל תהליך איטרטיבי הדומה ל-PR על מנת לחלק את האמון בין האתרים לפי מבנה הקישורים ביניהם.

ככל שמתרחקים מהאתרים הטובים שנבדקו ידנית, כך גדל הסיכוי שהאתר שאליו נגיע הינו אתר רע. נציין זאת באופן הבא: יהי דף  $p$  בעל אמון  $\theta$ , המקשר אל  $n$  דפים.

סה"כ האמון היכול להיות מועבר לדפי הבנים הינו  $\alpha_B \cdot \theta$ , כאשר  $\alpha_B$  הוא מקדם הדעיכה. בנוסף, האמון מתחלק באופן שווה בין כל אחד מהקישורים היוצאים מהדף. נניח כי לדף  $p$

יש 3 קישורים יוצאים, אזי כל אחד מהדפים יקבל  $\alpha_B \cdot \left(\frac{\theta}{3}\right)$  מהאמון של  $p$ .

בדומה לחישוב ה-PageRank נבצע עבור הדף אליו מקשרים סופרפוזיציה של כל ה-TrustRank הנכנס, וערך זה יהיה ערך ה-TrustRank המעשי שלו.

## 5. אלגוריתם TrustRank

לאחר שהצגנו את הרקע, נציג את אלגוריתם TrustRank, כפי שהוא מופיע במאמר המקורי.

### Function TrustRank

Input:

T - Transition matrix

N - Number of pages

L - Limit of oracle invocations

$\alpha_B$  - Decay factor for biased PageRank

$M_B$  - number of biased PageRank iterations

Output:

$t^*$  - TrustRank scores

Begin

```

// evaluate seed-desirability of pages
(1) s = SelectSeed(. . .)

// generate corresponding ordering
(2)  $\sigma$  = Rank({1, . . . ,N}, s)

// select good seeds
(3) d =  $0_N$ 
    for i = 1 to L do
        if O( $\sigma(i)$ ) == 1 then
            d( $\sigma(i)$ ) = 1

// normalize static score distribution vector
(4) d = d/|d|

// compute TrustRank scores
(5)  $t^*$  = d
    for i = 1 to  $M_B$  do
         $t^* = \alpha_B \cdot T \cdot t^* + (1 - \alpha_B) \cdot d$ 
    return  $t^*$ 
end
```

שלב 1 הינו בחירת קבוצת השורש. שלב 2-4 מגדירים את הערך לכל צומת כפי שתיארנו. סעיף 5 הינו השלב האיטרטיבי בו אנחנו מחשבים את TrustRank, ובסופו אנו פולטים את התוצאה.

## 6. בחירת קבוצת השורש

על מנת להפיק תוצאות טובות על האלגוריתם לבחור את קבוצת השורש באופן טוב, על מנת שהיא תהיה מייצגת וכן יהיה ממנה גישה אל רוב האתרים הטובים ברשת. יש לשמור על קבוצת השורש בגודל סביר, על מנת שקבוצת מומחים אנושיים אכן תוכל לעבור עליה ולדרג את האתרים בה.

כותבי האלגוריתם הציעו שלוש אסטרטגיות אפשריות לבחירת קבוצת השורש.

### 6.1 בחירת קבוצת שורש אקראית

האפשרות הראשונה והפשוטה ביותר היא לקחת קבוצה אקראית כלשהי של דפים ולקבוע שהם קבוצת השורש. זוהי אסטרטגיה פשוטה, ולא סביר שמנועי החיפוש משתמשים בה.

### 6.2 בחירת קבוצת שורש לפי PageRank

הגישה השנייה לבחירת קבוצת השורש הינה בחירת דפים בעלי PageRank גבוה. דפים בעלי PageRank גבוה הינם דפים אליהם מקשרים אתרים רבים אחרים. לפי הרעיון של אלגוריתם PageRank המקורי, ישנה סבירות גבוהה שאתרים אלו חשובים.

זוהי בהחלט אפשרות הגיונית היא לבחור את קבוצת הדפים הנ"ל:

- קיים סיכוי גבוה שאתר בעל PR גבוה הוא אתר טוב, שנרצה להכליל אותו בקבוצת השורש שלנו כמקור לאתרים טובים אחרים.
- במידה ואתר בעל PR גבוה הינו אתר ספאם, נרצה לדעת זאת ולסמן אותו ככזה.

חסרון בשיטה זו הוא האפשרות שלא כל האתרים ברשת האינטרנט נגישים דווקא מאחד מהאתרים בעלי ה-PR הגבוה, ולכן יתכן שנפספס אתרים טובים אחרים ברשת האינטרנט.

### 6.3. בחירת קבוצת השורש לפי PageRank הפוך

הגישה השלישית אומרת שנרצה לתת עדיפות לאתרים מהם ניתן להגיע אל אתרים רבים אחרים. ההגיון בגישה זו היא שעל ידי כך נוכל לבחור קבוצת שורש שתכסה שטח גדול ככל האפשר ברחבי רשת האינטרנט.

הבחירה נעשית על ידי חישוב של PageRank הפוך. החישוב נעשה בצורה דומה לחישוב ה-PageRank, רק שעבור כל דף, במקום להתייחס אל מספר הקישורים הנכנסים אליו, אנחנו מתרכזים במספר הקישורים היוצאים ממנו. ניתן למעשה להפוך את כיווני החצים עבור הגרף המקורי  $G$  וליצור גרף חדש  $G'$  עם הכיוונים ההפוכים. נחפש את הדפים ב- $G'$  בעלי ה-PR הגבוה ביותר, והם יהיו קבוצת השורש שלנו.

## 7. סיכום ומסקנות

האלגוריתם מציין נסיגה מהגישה של שוויון באינטרנט ונתינת הכוח לאנשים. מנועי החיפוש מרכזים כוח אצלם על מנת להגדיר את "שורש האינטרנט הטוב". מה יהיה שורש זה במידה וגוגל יישמו אצלם את האלגוריתם? קיים סיכוי טוב שאחד האתרים שיבחרו יהיה DMOZ. לגבי אתרים אחרים שייבחרו כשורש – זו שאלה מעניינת שאנשים רבים ינסו לגלות, אם וכאשר יאומץ האלגוריתם החדש.

סביר להניח גם שגוגל לא יחזרו על הטעות שביצעו עם אלגוריתם ה-PageRank, בכך שחשפו אותו לעולם ב-Google Toolbar, ויצרו למעשה תעשייה של סוחרים ב-PageRank. הגיוני שערך ה-TrustRank של האתרים ישמר הפעם אצל גוגל בסוד.

האלגוריתם החדש מעורר התלהבות בגוגל בגלל הקירבה שלו אל אלגוריתם ה-PageRank. מלבד התאמות קלות, אין להם צורך בכתיבת מערכת חדשה לחישוב ערך זה – ניתן להשתמש במערכת הקיימת של PageRank עם מספר התאמות קלות בלבד. כפי שאורן שץ ציין במאמרו בנושא, עם כל הרעש שמתחיל מסביב נושא ה-TrustRank, בגדול בסך הכל מדובר על גרסה חדשה ומשופצת של האלגוריתם הוותיק PageRank, המיועדת לצמצם את כמות תוצאות הספאם.

EOF