**Malicious Page Detector**

**דו" מסכם – מיני פרויקט במערכות הגנה לרשת:**

יובל עמרמי 316417591

יעל ינאי 316418656

רעיון הפרויקט:

תוסף לדפדפן chrome שיתריע בפני המשתמש על אתרי אינטרנט שחשודים כזדוניים. המטרה היא להגן על משתמשים מפני phishing – גניבת מידע רגיש מהמשתמש על ידי התחזות.

בעת גלישה באינטרנט, כאשר התוסף יופעל, אם יזוהה אתר חשוד, תיחסם בפני המשתמש את הגישה אליו. המשתמש יוכל להחליט לפי שיקול דעתו אם החסימה מוצדקת או לא, ועל פי החלטתו לאשר אותה או לאפשר את הכניסה לאתר.

מטרת הפרויקט:

המטרה המרכזית שלנו בפרויקט היא להגן על משתמש מפני כניסה ללינקים חשודים ברשת ולוודא שהקישור עליו לוחץ המשתמש מפנה לאתר תקין.

בנוסף, עוד מטרה שהייתה לנו בפרויקט הייתה להכיר טוב יותר את עולם התוכן הזה ולהבין את זרימת הנתונים ברשת בעת לחיצה על קישור ומעבר לאתר אחר.

תיאור יכולות התוסף:

התוסף שכתבנו נקרא – "Malicious Links Detector". נתאר את יכולותיו ואת דרך פעולתו:

* התוסף פועל ברקע מרגע פתיחת הדפדפן, ומופעל על כל אתר שהמשתמש ניגש אליו, ללא סינון או החרגה של אתרים ספציפיים.
* בעת לחיצה על התוסף תופיע רשימת כל כתובות ה-URL שמזוהות בעמוד בו הוא נמצא.
* הוגדרה רשימה של "blocked sites", האתרים שאליהם הכניסה חסומה. תוכן רשימה זו נשמר בין כניסות שונות לדפדפן ולא נמחק בכל סגירה שלו.
* בעת גלישה באתר, התוסף "סורק" את כל הקישורים המופיעים בעמוד בו נמצאים ומחפש אחר קישורים "זדוניים". ההגדרות שבחרנו להגדיר בעזרתן קישור זדוני הן אלו:

כתובות URL שאורכן קטן מ- 8 תווים, כתובות URL שאורכן גדול מ- 100 תווים, כתובות URL שמכילות בתוכן מילים חשודות (ההגדרה ל"חשודות" יכולה להשתנות על פי המשתמש).

המשתמש יכול לבחור להפעיל את התוסף על כל תת קבוצה של הגדרות אלו.

* ברגע שזוהתה כתובת חשודה על פי אחת מהסינונים שהמשתמש בחר להפעיל, קופצת חלונית על המסך שמציגה את הכתובות החשודות שהתוסף זיהה ושאלה למשתמש האם לחסום אותן. המשתמש יכול לבחור לחסום את כל הכתובות המוצעות, את חלקן או אף אחת מהן. הכתובות שיש לחסום יתווספו לרשימת הכתובות החשודות.
* למשתמש יש גישה לרשימת ה - blocked sites דרך הדפדפן.

בנוסף לצפייה ברשימה יש לו גם אפשרות לערוך אותה, וכך להוסיף בעצמו כתובות נוספות וגם להוריד כתובות שהוגדרו כחשודות על ידי התוסף או על ידי המשתמש עצמו בעבר.

* כאשר משתמש ינסה לגשת לאתר שנמצא ברשימת ה - blocked sites, בעת לחיצה על קישור שמפנה לאתר זה או פנייה לאתר דרך הקלדת הכתובת שלו בשורת ה – URL, הגישה תיחסם. כלומר, מה שיקרה בפועל הוא שייפתח tab חדש בדפדפן ומיד ייסגר, עוד לפני שהאתר ייטען.

פתיחת וסגירת ה- tab היא אינדיקציה למשתמש שהוא ניסה לגשת לאתר חסום.

* בנוסף לרשימת האתרים החסומים אליהם למשתמש יש גישה לעריכה, קיימת גם רשימה של מילים חשודות שאם התוסף יזהה שמופיעות בתוך כתובות URL יציע לחסום אותן. רשימה זו נבנית על ידי המשתמש, כאשר ההמלצה היא להכניס לרשימה מילים כגון prize, win וכו' שמופעיות בכתובת ה- URL עצמה (לדוגמא, www.winprize.com), מילים אשר יכולות "לפתות" אנשים ולשטות בהם כך שיחשבו שמדובר בלינק אמין ולכן יבחרו (שלא במודע) להיכנס לאתרים זדוניים.
* בזמן גלישה באתר אינטרנט, התוסף מתעדכן מחדש כל 5 דקות על הקישורים שיש באתר, כך ובמידה והקישורים משתנים הצעות התוסף יהיו עדכניות ומותאמות. בנוסף, בכל כניסה לחלון חדש התוסף מתעדכן ושואב שוב את נתוני העמוד.

מפרט טכני:

התוסף נכתב ב- JavaScript וב- HTML.

ב- JavaScript נכתבה הלוגיקה של התוסף – הגדרת פרמטרים לבדיקת תקינות של קישור וסינון אתרים שלא עונים על הפרמטרים הנדרשים. בנוסף בוצע שימוש בספריות של google chrome שנועדו לכתיבת תוספים ומאפשרים לגשת למשאבי הדפדפן. בכתיבת הקוד השתמשנו גם כן ב- jQuery.

ב- HTML וב- CSS הוגדרה הנראות של התוסף והוגדרה רשימת האתרים שהכניסה אליהם חסומה.

בעיות איתן התמודדנו:

הבעיה המרכזית שאיתה התמודדנו הייתה ההגדרה שעל פיה נפעל כדי לחסום או לאפשר גישה לאתר. בשלב זה התמודדנו עם שאלות כמו – מה מגדיר אתה חוקי/לא חוקי? איך ניתן לזהות על ידי הקישור לאתר או על ידי סריקה מהירה של אתר האם הוא זדוני? האם יש פרמטרים שנוכל להגדיר שיכריעו באופן מוחלט על חוקיות של אתר?

החלטנו בסופו של דבר להגדיר קריטריונים שעל פיהם האלגוריתם יקבע איזה אתר הוא חוקי ואיזה זדוני. הקריטריונים שבחרנו הם סינון על פי אורך כתובת ה- URL: כתובת שמכילה מתחת ל-8 תווים או מעל 100 תווים, וסינון על פי מילים "לא חוקיות" שמופיעות בתוך ה- URL.

מובן שיכולנו לבחור גם קריטריונים אחרים והתוסף היה פועל על פיהם באופן דומה לצורת פעולתו כרגע. בנוסף, ניתן תמיד להוסיף עוד קריטריונים ולהרחיב את אופן עבודת התוסף ואת יכולותיו.

בהקשר של קושי זה עלתה גם השאלה האם להגדיר קישור כזדוני על סמך ה- URL בלבד או על סמך נתונים נוספים. בתחילה חשבנו "לסרוק" את כל הטקסט שרלוונטי לקישור מסוים ועל סמך זה להתריע. אך פה עלתה בעייתיות, מסובך לחלק את הטקסט שמופיע בעמוד ולהתאים כל קישור לחלק אחר בטקסט, לעיתים יש חפיפה וקשה למצוא תבנית שתתאים לחלוקה שכזו. בנוסף, כדי להוסיף לקוד שכתבנו את יכולת זו נדרשו שינויים רבים ומעבר לעומק על כל רכיבי ה- HTML באתר ושינוי מתודולוגיה. שינוי כזה מאתגר גם מבחינת הטכנולוגיות בהן השתמשנו, לכן הוחלט בסופו של דבר להתמקד בלינקים שמכילים מילים שיכולות להיות חשודות, מילים שנועדו "למשוך" את הקורא ללחוץ על הקישור.

בעיה נוספת איתה התמודדנו הייתה תחזוקת רשימה של "אתרים טובים" כפי שהגדרנו בתוכנית העבודה של הפרויקט. לאחר בחינת האפשרויות והתחלת כתיבת המימוש הגענו למסקנה כי רשימה כזו של "אתרים טובים" תהיה מורכבת למימוש ולתחזוקה. זאת כיוון וככל שהמשתמש יגלוש יותר זמן באינטרנט הרשימה תלך ותגדל באופן לא פרופורציונלי (אנחנו מניחות במשפט זה שרוב האתרים בהם הציבור גולש חוקיים ואמינים). לכן, כדי להתמודד עם קושי זה הוחלט לשנות את הנחת הבסיס שלנו – במקום תחזוקה של רשימת אתרים מותרים לכניסה כך שכל אתר שלא ברשימה נחסמת הכניסה אליו, החלטנו להניח שכל אתר/ קישור שהמשתמש נתקל בו הוא חוקי, אלא אם כן יש סיבה לחשוב אחרת. כלומר, מימשנו רשימה של "אתרים רעים" ובכל פעם שהתוסף נתקל באתר שחשוד להיות ככזה הוא נכנס לרשימה.

עוד בעיה שעלתה הייתה שמירה על חווית המשתמש תוך כדי השימוש בתוסף. כדי לפתור זאת הגדרנו תוסף שפועל ברקע על כל אתר שנכנסים אליו ומתריע בצורה מסודרת למשתמש על כתובות חסומות שזיהה (לא מתריע על כל כתובת בנפרד, אלא מאגד רשימה של כתובות שנמצאו כחשודות לזדוניות ומציג למשתמש). למשתמש ניתנת האפשרות מתי שירצה להיכנס לרשימת האתרים ולצפות בה או לשנות אותה אבל אין צורך שזה יפגע בעבודתו.

חוזקות וחולשות של התוסף:

**חוזקות:**

* מתחזק רשימה של אתרים חסומים שניתנת לשינוי בכל רגע נתון (לא סופית וקבועה).
* לא פוגע בחוויית המשתמש בעת גלישה בדפדפן.
* פועל על פי לוגיקה מוגדרת מראש לשם חסימת אתרים.
* מאפשר למשתמש להוסיף ולהסיר אתרים בעצמו.
* מונע מהמשתמש כניסה לאתרים החסומים, ואף לא טוען אותם.
* זמינות – לא נדרש מהמשתמש לפתוח תוכנה או אתר נוסף במקביל לעבודתו במחשב, פועל על הדפדפן שבשימוש.
* הגנה על טעויות שנגרמות מחוסר ערנות של המשתמש (לחיצה על קישור ש"נראה אמין" והגיע כהודעת מייל, או מתוך אתר מוכר).
* הגנה על אוכלוסיות שמועדות יותר ל"נפילה" לטעויות כאלו (אנשים מבוגרים, ילדים...).
* עדכון המידע שהדפדפן פועל על פיו בכל טעינה של עמוד ובכל 5 דקות אם נשארים באותו העמוד.

**חולשות:**

* הלוגיקה שהגדרנו לא מכסה את כל המקרים (ניתן כמובן להרחיב אותה, אך עדיין לעולם נוכל לכסות את כל המקרים האפשריים).
* נקודת ההנחה שניתן לזהות אתרים זדוניים על פי קריטריונים לגבי כתובת ה- URL (אורך כתובת, כתובות שמכילה מילים מסוימות...)לא תמיד נכונה.
* ייתכן והתוסף יטעה ויגדיר אתר זדוני כתקין או להפך (false positive & false negative).

רכיבים מרכזיים בקוד:

* **קובץ manifest.json** – קובץ הגדרות. הצהרה על שם התוסף, גרסה, לוגו, הרשאות שיהיו לתוסף (בתוך דפדפן chrome, לאיזה יכולות תהיה לתוסף גישה), הגדרת metadata של התוסף וכו'.
* **קובץ background.js** – מכיל את הקוד שרץ ברקע ו"מאזין" לאירועים כל עוד התוסף פועל. מכיל לוגיקה פנימית של התוסף, את ה- state של האובייקטים שבו.
* **קובץ content.js** – רצים בהקשר של דפי web. האובייקטים שבקובץ מתקשרים עם ה- background. מאפשרים גישה ושינויים ב- DOM (Document Object Model) שבעמוד.
* **קובץ oprtions.html** – דף הגדרות של התוסף, יכולות להשתנות בזמן שהתוסף פועל.
* **קבצים options\_block.js, options\_block\_extra.js** – מכילים את הלוגיקה של הכפתורים ואפשרויות הבחירה שיש בדף options.
* **קובץ popup.js** – ה- UI שרואים כשלוחצים על הלוגו של התוסף בדפדפן.

\* תיאור של הפונקציות המרכזיות בהן השתמשנו נמצא בתיעוד שבתוך הקוד עצמו.

מענה הפרויקט למטרה הראשונית:

אנחנו מאמינות שהפרויקט אליו הגענו עונה על המטרות המרכזיות אותן הגדרנו לפני תחילת כתיבת המימוש עצמו. התוצר המוגמר הוא תוסף לדפדפן chrome שאכן חוסם אתרים שחשודים כזדוניים ומונע כניסה אליהם, אלא אם כן המשתמש בצורה אקטיבית בוחר לאפשר את הכניסה אליהם.

תוך כדי כתיבת קוד התוסף השתנו חלק מהנחות הבסיס שלנו, אך ההנחות החדשות נראות לנו נכונות ומותאמות יותר בעיה אותנו אנו רוצות לפתור. לדוגמא, תחזוקה של רשימת "אתרים טובים" (יעד ראשוני שהגדרנו) לעומת תחזוקה רק של רשימת "אתרים רעים" (מה שנעשה בפועל).

כפי שהגדרנו בתוכנית העבודה, הגענו למוצר שמזהה כתובות חשודות, מתריע למשתמש ומאפשר כך חוויית גלישה באינטרנט בטוחה יותר תוך מזעור סיכויי כניסה לאתרים שמבצעים phishing.

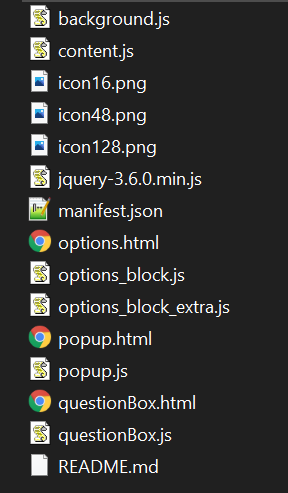
סיכום:

התוסף אותו כתבנו נוח מאוד לשימוש ולתפעול ולא דורש השקעת זמן בלמידתו, הוא מותאם לכלל האוכלוסייה שגולשת באינטרנט.

בחרנו ליישם לוגיקה בסיסית להגדרת כתובת כחשודה להיות זדונית, אך זאת רק לשם הוכחת היתכנות (POC), מובן כי ניתן להרחיב אותה ובכך להרחיב את טווח ההגנה של התוסף. רעיון לפיתוח עתידי שעליו חשבנו תוך כדי מימוש הפרויקט הוא הוספה של אלגוריתם לומד שיבנה דפוסים מהאתרים בהם המשתמש גולש ומהקישורים עליהם הוא לוחץ וגם ילמד מהרשת על ידי חיפוש של כתובות לא מוכרות ויצירת מנגנון רחב יותר לזיהוי של כתובות זדוניות.

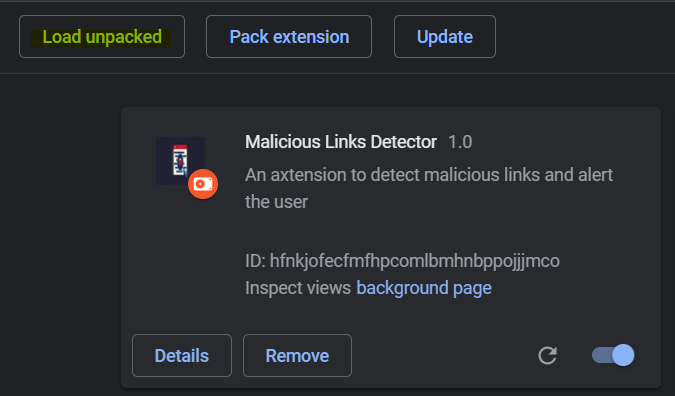
במהלך כתיבת הפרויקט התעסקנו בנושאים אותם לא הכרנו לעומק לפני כן, דנו רבות ברעיונות מימוש של התוסף ובכך חקרנו מנגנוני זרימת נתונים ברשת וגם לוגיקות שונות שניתן להגדיר על מנת לסנן מידע או על מנת להגדיר כתובות URL כתקינות או שלא. למדנו הרבה על עולם תוכן זה ובנוסף למדנו גם כיצד לכתוב נכון תוסף לדפדפן chrome והתעסקנו בטכנולוגיות חדשות בהן לא עסקנו לפני כן.

התקנת התוסף:

כדי להתחיל להשתמש בתוסף יש להוריד למחשב ולשמור מקומית בתיקייה את הקבצים הבאים:

בשלב הבא יש להיכנס בדפדפן ה- chrome ל- chrome://extensions.

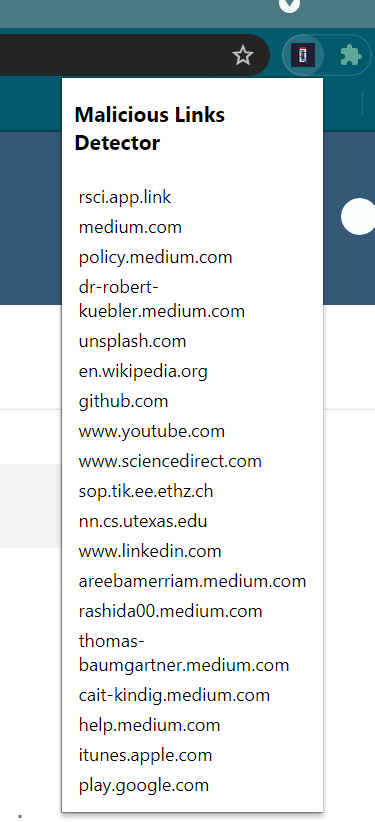
משם יש לטעון את התוסף החדש. לוחצים על הכפתור Load unpacked ובוחרים בתיקייה שבה נמצא הקוד לתוסף. לאחר מכן התוסף יופיע בעמוד, יחד עם שאר התוספים שהותקנו ב- chrome.



כעת התוסף מותקן וניתן להתחיל להשתמש בו (דוגמת שימוש מפורטת מטה).

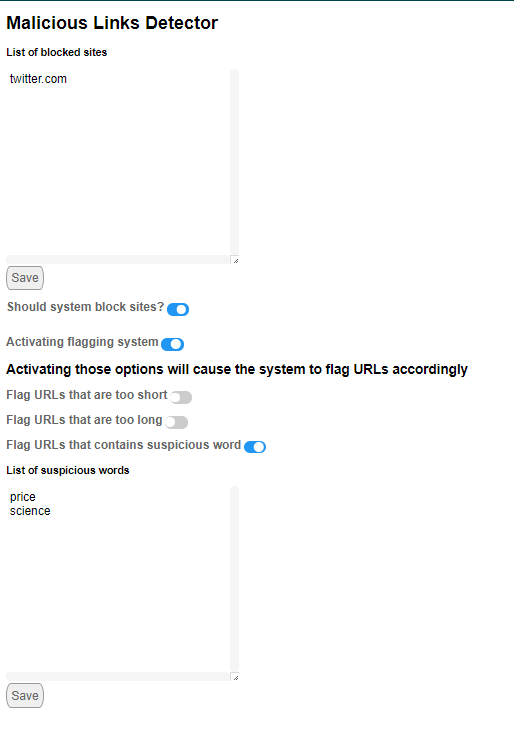
כדי לעצור את פעולתו יש ללחוץ על הכפתור שמופיע בתמונה מצד ימין למטה. ניתן לעשות זאת בכל שלה

דוגמת שימוש בתוסף:

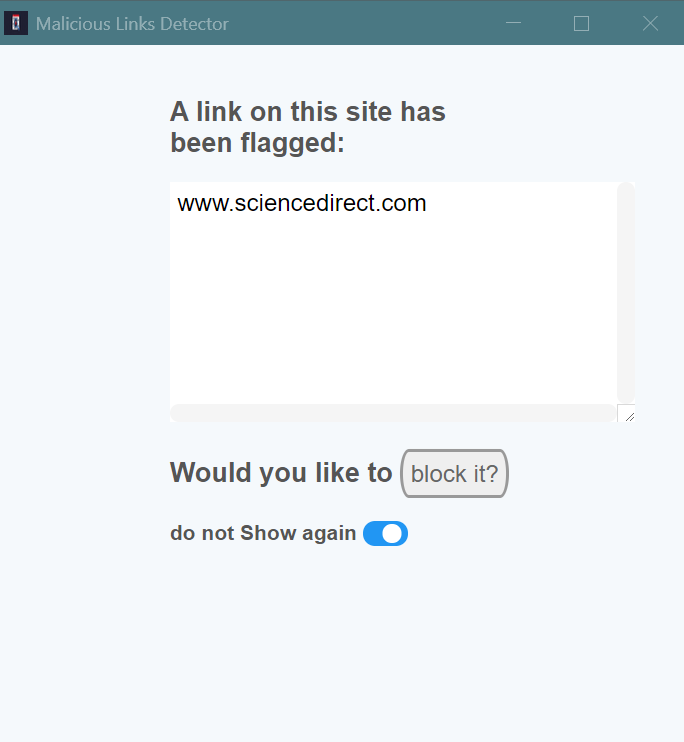
בעת כניסה לאתר כאשר התוסף דולק ולחיצה על הלוגו שלו תופיע רשימת כל הקישורים החיצוניים שיש בעמוד הנוכחי בו המשתמש נמצא בתוך האתר.

לחצן ימני על הלוגו של התוסף ובחירה ב- options תעביר את המשתמש לדף הבא בו הוא יוכל להגדיר את התוסף לצרכיו ולרצונו.

בחלק העליון נמצאת רשימת האתרים החסומים. לכאן המשתמש יכול להכניס אתרים באופן ידני ובנוסף בכל פעם שיקבל הצעת חסימה של התוסף הכתובת שאישר תתווסף לרשימה.

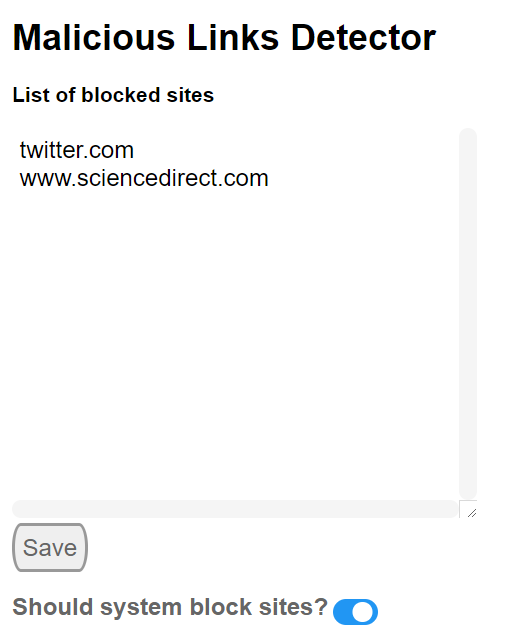
בחלק השני של העמוד ניתן לראות את ה- flags השונים שהמשתמש יכול להפעיל עבור חסימה. ניתן לשנות את הבחירה בכל רגע נתון.

בהינתן שמשתמש הפעיל את ה- flag שמאתר בתוך ה-URL מילים שהוגדרו כחשודות, בחלק התחתון של העמוד הוא יכול להוסיף לרשימה את המילים שיבחר והן יישמרו בהגדרות הדפדפן.

נניח ומשתמש הכניס לרשימת המילים החשודות את המילים שמופיעות בתמונה למעלה.

בזמן גלישה באתר אם אחת מהמלים ברשימה זוהתה בתוך URL, תקפוץ החלונית הבאה למסך ובה כל הכתובות שלא עברו את הסינון. המשתמש יכול ללחוץ על ה- X ולא לאשר אף כתובת, למחוק חלק מההצעות שברשימה, או לאשר את כולן. כדי לאשר ילחץ על block it.

בנוסף, אם משתמש הפעיל את ההתרעות גם אל אורכי הכתובות וזוהו כתובות חשודות, גם הן יופיעו ברשימה זו שמופיעה על המסך.



נניח ומשתמש בחר לאשר את ההצעה שמופיעה למעלה, כעת בכניסה לעמוד options של התוסף תתווסף הכתובת החדשה, וכל עוד היא ברשימה כניסה אליה לא תתאפשר.