מעבדת סייבר - 1

מטלת תכנות זיהוי כופרה

בתרגיל זה ננסה לבנות כלי המגן נגד התקפות כופר.

הכלי אמור להיות מורץ על תיקיה מסוימת, ולוודא שאכן אף קובץ לא הוצפן או שונה בדרך שלא רצויה. הכלי עובד בסביבת עבודה אנושית נורמטיבית בה אנשים עובדים עם הקבצים בתיקייה. כלומר, הקבצים במערכת יכולים לעבור שינויים תקינים לחלוטין.

בתיקיה ימצאו קבצי טקסט(.txt) המכילים תוי ascii בלבד. הקוד שלכם יכתב בשפת פייטון, בזוגות.

אתם יכולים לצאת מנקודת הנחה כי ההצפנה לא בהכרח מתבצעת על כל קבצי התיקיה יחד, או על קובץ שלם בהכרח. בנוסף, ההצפנה אינה בהכרח, הופכת את תוכן הקובץ ל-"ג'יבריש".

כמו כן, אתם צריכים לייצר כלי שיהיה יעיל מבחינת משאבים.

הצעה לפתרון - עליכם לחשוב איך תבדקו שקובץ מסוים השתנה - ומשם, האם השינוי הינו חוקי. יכולות להיות מספר גישות, בחרו את זו המתאימה לכם. **שימו לב לדגשים, הם חלק מדרישות המטלה.**

בונוס (לציון הקורס) יינתן לתוכנות זיהוי יוצאות דופן, מבחינת היצירתיות, והיעילות החישובית.

**דגשים:**

1. במידה ולקחתם שיטה ידועה, מתוך מאמר/בלוג וכו', יש לציין זאת בהערת קוד בתחילת הסקריפט שלכם. אי אזכור של דבר שכזה יכול לפסול את המטלה בתור העתקה. כמובן, העתקה של פתרון מקבוצה אחרת יגרום לפסילה של המטלה.

2. יש לנתח את *יעילות המערכת* לפי שלושה היבטים עיקריים:

א. \*\*שימוש בזיכרון\*\* – כמה מידע נשמר בזיכרון? לדוגמה: האם אתם שומרים עותק של כל הקבצים או רק חלק?

- ערכים נורמטיביים:

- ((O(1): אם רק נתוני מטא (timestamps) נשמרים

- (O(n)): אם נשמרים hash-ים של כל קובץ

- (O(n\*m)): אם נשמרת גרסה מלאה של תוכן (כאשר m הוא גודל ממוצע של קובץ)

ב. \*\*זמן ריצה\*\* – מה זמן הריצה של סריקה אחת או של גילוי הצפנה?

- ערכים נורמטיביים:

- (O(n)): סריקה מלאה של הקבצים

- ((O(log(n))): אם עוברים על המידע בצורה של עץ.

ג. \*\*סיבוכיות קלט/פלט (I/O)\*\* – כמה קריאות דיסק מתבצעות לצורך זיהוי שינויים?

- ערכים נורמטיביים:

- נמוך: שימוש באירועים מערכתיים

- גבוה: סריקה תקופתית (polling) כל X שניות

את הניתוח יש להוסיף בתוספת הסבר על כל היבט בהערת קוד בתחילת הקובץ.

3. את המטלה יש להגיש לתיבת ההגשה של "זיהוי כופרה" , עם פורמט של ת.ז. עוקבות בקובץ זיפ. לדוגמא, אם שני המגישים הם ת.ז. 3333333333 ו999999999, אז הקובץ הוא 333333333\_999999999.zip.