XX. Fiatal Műszakiak Tudományos Ülésszaka, 2015. Kolozsvár, 147–150. *DOI: 10.33895/mtk-2015.03.29*
http://hdl.handle.net/10598/28657

FELHASZNÁLÓI VISELKEDÉSANALÍZIS EGY LEHETSÉGES MÓDSZERE

REAL-TIME USER BEHAVIOR ANALYSIS - MACHINE LEARNING IN IDENTITY INTELLIGENCE

Fried Zoltán

BalaBit IT Security, Cím: 1117, Magyarország, Budapest Alíz utca 2.; Telefon: +36 1 398 6700, Fax: +36 1 208 0875, zoltan.fried@balabit.com

Abstract

A real-time user behavior analytics solution that mitigates the impact of advanced persistent threats (APTs) and potential data breaches helps.

***Keywords**: user behavior, secirty, analysis, real-time.*

Összefoglalás

Napjaink IT-biztonságtechnikai és erre épülő riasztórendszerei (SIEM) a pontról pontra definiált egymásra épülő szabályrendszereken alapulnak. A rohamosan fejlődő és növekvő számítógépes rendszerek (pl. felhők) és még nagyobb számú hozzájuk köthető biztonságtechnikai szabályrendszerek már szinte kezelhetetlen terhet jelentenek ezen rendszereket üzemeltető szakemberekre. Ezen rendszerek beüzemelésének és üzemeltetésének erőforrásigényén enyhít a dolgozatban ismertetett öntanuláson alapuló, valósidejű, új generációs felügyeletei technológia.

***Kulcsszavak:** viselkedésanalízis, biztonság, riasztás, felügyelet, valósidejű.*

1. Bevezetés

A számítógépes rendszerek felhasználói jogosultságainak a beállítása és karbantartása több száz vagy ezer felhasználó esetén nehéz és bonyolult feladat. Az egyes felhasználókhoz köthető egyedi jogosultságok naprakész tartása, és a jogosultság-szegési kísérleteinek feldolgozása még nehezebb. A dolgozatban ismertetünk egy merőben új eljárást a fenti probléma egyszerű megvalósítására.

2. A jelenleg használt megoldások előnyei és hátrányai

A vállalatok az adataik védelme érdekében egyre több korlátozó szabályt hoznak és informatikai megoldásokat vezetnek be. Ezek a korlátozások az adatok védelme mellett a munkafolyamatokat gördülékenységét is korlátozzák, ami a termelékenységre is közvetlen hatással van. Ha egy felhasználó az informatikai rendszerben megkísérel megszegni egy szabályt, arról értesítés keletkezik, amire meghatározott időn belül reagálni kell. Ez addig nem is gond, amíg ezek az értesítések, riasztások száma viszonylag alacsony. Egy riasztás értéke attól is függ, hogy, milyen közel van a reakció az esemény bekövetkezési idejéhez. A reakció időt csak akkor lehet alacsonyan tartani, ha a nagy számú véletlenül kiváltott riasztások számát is alacsony értéken tartjuk és elsősorban csak azokkal a riasztások-

kal foglalkozunk amik valós veszélyt jelentenek a vállalat adatira.

A jelenlegi biztonságtechnikai megoldások használatakor pontosan meg kell mondani, hogy egy felhasználó mit tehet és mit nem egy informatikai rendszerben. Egy közepes vagy nagyvállalat esetében egy felhasználónak a munkájához általában több informatikai rendszerhez hozzá kell férnie. Ezek között biztos van olyan rendszer is ami valamilyen elvek alapján a vállalat számára érzékeny adatokkal dolgozik. Minél több párhuzamos informatikai rendszerhez fér hozzá egy felhasználó, annál fontosabb a vállalat számára, hogy tudja, hogy az adott felhasználó csak a neki szükséges erőforrásokat használja, és a munkája során megszerzett ismeretekkel nem él vissza.

A napjainkban működő informatikai rendszerek biztonságtechnikai megoldásai azon alapulnak, hogy a szükséges erőforrásokhoz hozzáférési jogosultságot adnak, minden egyéb erőforráshoz viszont a hozzáférést tiltják. Az informatikai szakemberek a karbantarthatóság figyelembevételével nagy számú felhasználó esetében a jogosultságokat és a felhasználókat külön csoportokba rendezik, és ezeket a felhasználó csoportokat és a jogosultság csoportokat rendelik egymáshoz. Nagyon ritka eset, amikor egy valamilyen szempont alapján kitüntetett felhasználó teljesen egyedi jogosultságokat kap - ezt minden eszközzel kerülni kell. A jogosultsági rendszer erőssége a pontos jogosultságspecifikáció, és az ebből fakadó jól körülhatárolható jogosultsági mátrix, hátránya viszont a nehéz karbantarthatóság, a szabályszegésekből (szándékos vagy véletlen) adódó nagy számú riasztások priorizálása és kezelése. A vállalati folyamatok informatikai leképzésének bonyolultságából adódóan a SIEM (Security information and event management) rendszereket nagyon nehéz és időigényes feladat megfelelően beállítani, a változó igényekre folyamatosan naprakészen tartani.

3. A felhasználói viselkedés-analízis

A felhasználói viselkedésanalízis egy viszonylag új terület az informatikai biztonságtechnika területén. Az elmélet azon alapszik, hogy ha ismerjük a vállalatban dolgozó felhasználók munkája során keletkezett tevékenységeket, akkor e tevékenységek időben folyamatos vizsgálata során minden egyes felhasználóra egyedileg jellemző ujjlenyomat képezhető. Ezen ujjlenyomat alapján megmondható, egy számítógépes rendszerben végzett tevékenységről, hogy a beleillik e a tevékenységet végző felhasználó profiljába vagy sem. Ha nem illik bele az előre meghatározott profilba, akkor kiszámolható, hogy ez mennyire tér el az ugyanazt vagy hasonló feladatot ellátó felhasználók viselkedésétől. A kapott eredmények alapján könnyen rangsorolhatóak az egyes események, riasztások.

4. Az új rendszer

4.1. A felhasználói viselkedés-analízishez szükséges profil kialakítása

Vegyük sorra egy felhasználó napi tevékenységét. Megérkezik a munkahelyére és bejelentkezik a számítógépes rendszerbe. Általában egy bejelentkezéssel több (SSO) informatikai rendszer használatához is jogosultságot szerez. A felhasználó a munkája során általában leveleket ír, dokumentumokat szerkeszt, a képernyőn az felhasználó számára aktív elemekre kattint, meghatározott helyekre jól körülhatárolható válaszokat gépel, dokumentumokat, cikkeket olvas, képeket, videókat néz vagy esetleg szöveges parancsokat ad a számítógépnek. Ezek az információk még akkor is leképezhetőek a felhasználó számítógépétől független adatbázisokba, ha arról a felhasznál nem adott kifejezett parancsot. Abban az esetben ha ezeket az információmorzsákat összegyűjtjük, akkor képesek vagyunk valós idő-

ben reagálni a felhasználó által generált eseményekre.

4.2 Az adatforrások

- A 4.1-es fejezetben említett információforrások közül a legelterjedtebb a login/logout időpont - felhasználó - informatikai rendszer eseményhármas. Ezen eseményhármas kiegészíthető egy olyan információval is, hogy milyen rendszeren keresztül végezzük a be vagy kijelentkezést;

- Az IT szakemberek által kiadott parancsok. Ezek a parancsok mintázata nagyon jellemző a parancsot kiadó szakemberre, még akkor is, ha több szakembernek ugyanaz a feladata, ugyanazokat a parancsokat adják ki;

- A felhasználó a munkája során milyen szervereken dolgozik, ezt mikor teszi és ott milyen feladatot hajt végre. Ebben a pontban tulajdonképpen minden információ benne van, ami a feladatra jellemző, kivéve azt az információhalmazt amin a feladatot el kell végezni.

4.3 Az algoritmusok egyszerűsített működése

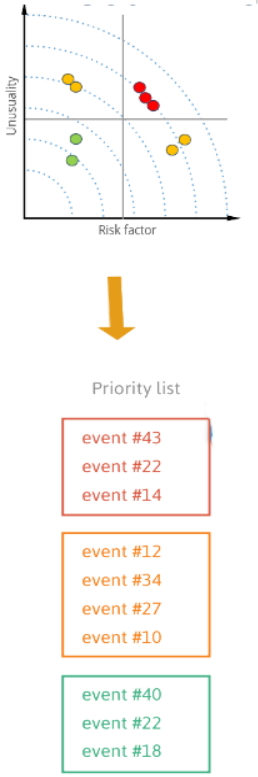
A be- és kijelentkezési információk vizsgálatához a normál eloszlás használjuk. Leegyszerűsítve ezeket a valószínűségeket összegezzük és a eredmény képezi a felhasználó profilját. Ennek a vizsgálatnak az a célja, hogy meg tudjuk határozni, hogy a felhasználó a következő be- és kijelentkezésének mekkora a valószínűsége a korábbi eredményekhez képest. Ha ez az eltérés átlép egy küszöböt riasztás generálódik a rendszerben. Ha megkülönböztetjük a sikeres és sikertelen kísérleteket és azok számosságát is figyelembe vesszük, akkor következtetni lehet az egyes behatolási kísérlet jogosultságának a fokára is.

A vásárlói szokások vizsgálatához használt recommandation enginek [1] módosított változati nagyon jól felhasználhatóak ehhez a területhez is. Ilyenkor az engine által felépített mátrixot arra használjuk, hogy az adott pillanatban bekövetkezett esemény valószínűsége mennyivel tér el az engine által jósolttól. Ehhez a vizsgálathoz a mátrix faktorizációt [2] használjuk. A mátrix felépítéséhez az egyes dimenziók "szöveges" megfogalmazását számszerűsíteni kell, majd addig transzformáljuk az adatokat, amíg a mátrix egyes dimenziói ortogonálisak lesznek egymásra. Riasztásra akkor kerül sor, amikor a mátrix mindegyik dimenziójához rendelt riasztási küszöböt átlépte az esemény bekövetkezésének a valószínűsége.

Egy másik lehetséges eljárás a vásárlói kosár elemzéséhez használt algoritmusok [3]. Ilyenkor azt vizsgáljuk, hogy a felhasználó az egyes hozzáférési kísérletek során, vagy a számítógépnek kiadott utasításai pillanatában a művelethez szükséges adatok mellett az egyes járulékos paraméterekből (például: kliens IP cím, szerver IP cím, protokollok, portok, autentikációs paraméterek, alkalmazások, stb.) számított pontérték milyen távol helyezkedik el a felhasználó korábban felépített profiljában eltárolt pontéréktől. A távolság függvényében küldünk riasztást az éppen folyamatban levő tevékenységekről.

5. Következtetések

Az IT-biztonságtechinka evolúciójának a legújabb generációjához tartozó megoldások az informatikai rendszerekből származó információk összegyűjtésével és elemzésével foglalkoznak. Ezek olyan vizsgálati módszereket alkalmaznak, melyek valós idejű, átfogó képet tudnak mutatni a rendszerben éppen történő eseményekről, amelyek segítségével az IT-biztonságtechnikai szakemberek a valós kockázatokat jelentő eseményekre koncentrálhatnak.



1. ábra. *A rendszerbe érkező események kategorizálása és érzékenységük szerint sorrendbe állítása*

A technológia lényege az, hogy nem egy újabb és még szigorúbb irányelveken alapuló kontrollmegoldások bevezetését célozza meg, hanem a felhasználók egyedi viselkedésmintáinak monitorozásán, az aktivitásokban, a szokásostól eltérő anomáliák matematikai algoritmusokkal történő kiszűrésén és azok fókuszált kivizsgálásán alapul. A tapasztalatok alapján a folyamatos, valós idejű tevékenységmonitorozás előnye az, hogy a vállalati rendszerek biztonságát úgy képes növeli, hogy közben sem az üzleti folyamatokat, sem a felhasználók mindennapi munkavégzését nem korlátozza.

Szakirodalmi hivatkozások

[1] Barry K. Lavine: *Clustering and Classification of Analytical Data*, Clarkson University, Potsdam, USA, Encyclopedia of Analytical Chemistry, ISBN 0471 97670 9

[2] Ruslan Salakhutdinov and Andriy Mnih: *Probabilistic Matrix Factorization*, Department of Computer Science, University of Toronto, 6 King's College Rd, M5S 3G4, Canada

[3] Renáta Iváncsy and István Vajk: *A time- and memory-efficient frequent itemset discovering algorithm for association rule mining*, Int. J. Computer Applications in Technology, Vol. 27, No. 4, 2006, 270-280.