HADMÉRNÖK

IV. Évfolyam 3. szám - 2009. szeptember

Gyányi Sándor
gyanyi.sandor@kvk.bmf.hu

BOTNETEK FELKUTATÁSA A TÉRINFORMATIKA SEGÍTSÉGÉVEL

Absztrakt

Világunkban az információ egyre nagyobb szerepet kap, immár mindennapi életünket is nehezen tudjuk elképzelni informatikai rendszerek használata nélkül. Interneten bármit el tudunk intézni, legyen az szállásfoglalás a világ másik oldalán, pénzátutalás vagy akár pizzarendelés. A nagyobb szerep nagyobb felelősséget is jelent minden felhasználó számára, hiszen esetenként komoly összegek úszhatnak el a nem kellő gondosság miatt. A gombamód szaporodó rosszindulatú alkalmazások veszélyeztethetik az emberek információs infrastruktúrába vetett bizalmát. A korszerű módszereket használó bűnözők általában más, ártatlan felhasználók álcája mögé bújnak, így utólagos azonosításuk meglehetősen nehéz. Mint sok más káros folyamat esetében, a számítógépes támadások esetében is hatásosabb a megelőzés, az Internet struktúrája azonban megnehezíti az egyes végpontok földrajzi elhelyezkedésének felderítését. Ebben a cikkben igyekeztem megvizsgálni azokat a lehetőségeket, amikkel automatikusan elvégezhető egy számítógépes végpont földrajzi koordinátákhoz rendelése. Egy ingyenes – és ennek megfelelően kevésbé precíz – adatbázis segítségével egyszerű térinformatikai alkalmazást készítettem, amivel fertőzött hálózati végpontokat, úgynevezett botnet tagokat lehet gyorsan és könnyen áttekinthető térképen ábrázolni. A végpontok azonosításához a botnetek egyik tevékenységének - nevezetesen a kéretlen levelek küldésének – eredményeit használtam fel.

Nowadays information technology is playing more and more important part and our everyday life can be hardly imagined without using it. We can arrange anything on the Internet, either hotel booking in the other side of the World, transferring money or ordering pizza. Wide range of possibilities entails more responsibility of the user, since there can be a lot of money lost because of the lack of carefulness. Trust in information infrastructure can be endangered by fast multiplying malicious applications. Criminals using modern technology are generally hiding in masks of innocent users, so it is very hard. to identify them subsequently. As in many other damage cases, prevention is the most effective way at computer attacks. However, the structure of the Internet makes it harder to find the geographical location of each hosts.

In my article I have tried to examine the possibilities with which every host can be automatically linked to some geographical coordinates. By using a free (and therefore not so precise) database I have made a simple GIS application, and with its help the infected net hosts (so-called botnet members) can be presented on a clear map. For identification of the hosts I used the results of sending unwanted e-mails - as one of the activities of botnets.

Kulcsszavak: *Botnet, SPAM, kéretlen levél, GeoIP, Google Maps*

Kéretlen levelek küldése botnetek segítségével

A botnetek napjaink egyik legveszélyesebb hálózati eszközei. Különleges, kifejezetten erre a célra készített számítógépes alkalmazások, amelyek a felhasználó tudta és engedélye nélkül kerülnek a számítógépre. Az ilyen megfertőzött gépek központi irányítás alá kerülnek, azt a feladatot hajtják végre, amit a gazdájuk (akit szokás „botmaster" vagy „botherder" néven is emlegetni) ír elő nekik. Több ilyen fertőzött, központilag irányított számítógép hálózatot alkot, amelynek szokásos neve a „botnet". Egy ilyen botnet földrajzilag elosztott rendszer, a tagok nagy száma miatt komoly erőforrásokkal rendelkezik, így szinte bármilyen, nagy számítási kapacitást vagy hálózati sávszélességet igénylő feladatra alkalmas. A kialakított botnetek alkalmazási területei közül említést érdemel a túlterheléses támadások (DDoS) kivitelezése, az adatlopás, csalások elkövetése illetve a kéretlen reklámlevelek küldése, amelyből a botmasternek bevétele is származik. A kéretlen reklámok kiküldése közben a botnet tagjai saját maguk is megpróbálhatnak újabb tagokat „toborozni", amihez kiváló lehetőség egy megfelelően fertőzött elektronikus levél. A kéretlen elektronikus levelek aránya napjainkban jóval nagyobb, mint a normál leveleké.

A kéretlen levelek – SPAM – küldése több fontos fejlődési szakaszon esett át. Kezdetben az ilyen levelek küldői saját üzemeltetésű, általában valamilyen egzotikus – és így nehezen lenyomozható – levelező szerverről indították kampányaikat. Ezek ellen hamarosan megjelentek a védekezési módszerek, egyszerűen a címzettek postafiókjait kezelő szerverek nem fogadtak el leveleket az ilyen, ismertté vált IP című szerverektől. A spammerek következő módszere az álcázás volt. Igyekeztek felkutatni olyan, tőlük teljesen függetlenül működő szervereket, amelyek rendszergazdája nem volt elég gondos és helytelenül konfigurálta fel rendszerét. A levelek továbbításában részt vevő, ún. SMTP szerverek ugyanis képesek lehetnek továbbítási (relay) funkciókra, vagyis átvehetnek olyan elektronikus leveleket, amelyek címzettje nem a saját postafiókok között található. Az átvétel után a feladó nevében továbbítják a tényleges címzettnek. Megjegyzendő, hogy az internet-szolgáltatók által az előfizetőiknek biztosított SMTP szolgáltatás is ilyen módon működik. A szolgáltatók saját előfizetőiket meg tudják különböztetni a többi internetezőtől, így a levéltovábbítást csak ennek a zárt csoportnak engedélyezik. Az olyan levelezőszervereket, amelyek bárkitől elfogadnak leveleket és azt bármely, nem saját kezelésű email címre továbbítják, „Open Relay" néven szokás említeni. A védekezés ezek ellen már nehezebb, hiszen ilyen szerver sok van, ráadásul ezeket a tulajdonosaik nem csak kéretlen levél küldésére használják. Több, ilyen Open Relay szerverek IP címeit tartalmazó adatbázis jött létre, amelyeket a levelező szerverek figyelhetnek, és a tiltólistában szereplő szerverektől visszautasíthatják a levél elfogadását.

A következő fejlődési lépcső olyan hálózati végpontok keresése volt, amelyeket a kéretlen levél küldője saját céljaira fel tud használni. A botnetek minderre egyszerű, és hatalmas teljesítményű megoldást kínálnak. Mivel a levelezőszerver is csak egy számítógép (amelyen SMTP alkalmazás fut), így egy egyszerű otthoni PC is alkalmas ilyen feladatokra. A botnet számos tagját felhasználva lehetővé válik egy időben, egyszerre sok végpontról indítani a

levelek küldését. Az otthoni PC-k által használt IP címek nem szerepelnek a feketelistákban, ráadásul, ha némelyiket kitiltják, másik lép a helyére. Nem elhanyagolható problémát jelent a legtöbb, háztartásokban használt internet előfizetés dinamikus IP cím kiosztása is. A szolgáltatók igyekeznek spórolni a nyilvános IP címek kiosztásával, mivel az ilyen címeket meg kell vásárolniuk. A szolgáltatók összes ügyfele egyszerre valószínűleg sosem csatlakozik a hálózatra, így felesleges mindegyiknek ilyen „drága", egyedi címet biztosítani. A gyakorlat az, hogy a hálózatra csatlakozás során a szolgáltató az éppen nem használt címekből ad egyet az ügyfélnek, folyamatos csatlakozás esetén pedig általában 24 óránként újítják meg a címet. Emiatt egy átlagos végpont IP címe 24 óránként megváltozik, vagyis ha egy bot kliens IP címe tiltólistára kerül, akkor 24 órán belül egy újabb címmel szabadon folytathatja a levélküldést.
Napjainkban botnetek felelnek a káros levelek küldésének túlnyomó többségéért. 2008-ban ez az arány 90% feletti volt [9].

A káros levelek szűrésének lehetőségei

A kéretlen reklámlevelek és vírusos csatolást tartalmazó levelek elleni védekezés nem egyszerű. Egy elektronikus levél elküldése a következő lépcsőket tartalmazza (abban az esetben, ha a küldő nem egy továbbító szervert vesz igénybe):

1. A levelet küldő (kliens) a címzett email címe alapján lekéri a domain névhez (az email cím „@" utáni része) tartozó DNS MX (Mail Exchanger) rekordot. Ez a rekord tartalmazza a domain névhez tartozó levelezőszerver elérhetőségét.
2. Az MX rekord alapján meghatározza a címzett levelezőszerver IP címét.
3. A kliens TCP kapcsolatot hoz létre a címzett szerver 25-ös portjára.
4. Az SMTP használatával megadja a feladó és a címzett email címét.
5. Elküldi a levelet.
6. Bontja a TCP kapcsolatot.

Mint látható, mindez nem igényel komolyabb erőforrásokat, egy ilyen kommunikációt lebonyolító SMTP „motor" néhány tíz kilobyte méretben megvalósítható, így egy bot kliensbe is beépíthető. A fogadó szerver a 3. pont után hajthatja végre az első ellenőrzési folyamatot. Mivel a TCP kapcsolatban a feladó IP címe nem, vagy csak nagyon nehezen hamisítható, így a küldő kénytelen a tényleges IP címét használni. A fogadó szerver végrehajthat néhány biztonsági funkciót:

- Ellenőrizheti, hogy az IP cím szerepel-e a tiltólistában.
- Ellenőrizheti, hogy a küldő SMTP szerver IP címéhez tartozó fordított DNS (Reverse DNS) adat megegyezik-e azzal, amit a szerver magáról hirdet (az otthoni számítógépek által használt IP címekhez nem mindig állít be a szolgáltató Reverse értéket). Ha nem egyezik meg, akkor a küldőtől nem fogadja el a levelet.

A leghatékonyabb módszer a folyamat ezen fázisában elutasítani a levél átvételét, hiszen ekkor történik a legkevesebb művelet elvégzése. A következő szűrési lehetőség a 4. pontban van, vagyis a feladó és a címzett email címének megadásakor. Ekkor a fogadó szerver:

- Visszautasíthatja a levéltovábbítást (ha a címzett email címe nem a saját postafiókok között található).
- Ellenőrizheti a feladó címét. Ha elindít egy levélküldést a feladó részére, és a címet kezelő szerver ismeretlennek jelzi vissza, akkor a feladó címe nem létezik, így a levél kéretlennek tekintendő. Ez egy erőforrás igényes folyamat, ráadásul a mostani kéretlen leveleket többnyire létező (a tulajdonos tudta és engedélye nélkül használt) email címet megadva küldik, tehát a hatékonysága is alacsony.
- A feladó címét keresheti tiltólistán. Ha szerepel, akkor a levelet nem veszi át.

Az utolsó ellenőrzési fázis a levél átvétele után lehetséges, ekkor már a levél teljes terjedelemben rendelkezésre áll, vagyis annak tartalma megvizsgálható. A tartalomvizsgálatra

több, szofisztikált módszer is létezik. Egyszerűbb esetben csak bizonyos kulcsszavak meglétét ellenőrzik, de lehetséges a valószínűségszámításra alapozott, öntanuló rendszereket is használni. A Bayes-tételre alapozott Bayes analizátor segítségével nem egyetlen szó előfordulását keresik, hanem a korábban érkezett levelekben található szavak együttes előfordulási valószínűségét vizsgálják. Ha bizonyos szavak kéretlen levelekben gyakrabban fordulnak elő együtt, és egy vizsgálandó levélben ugyanezek a szavak megtalálhatók, akkor a levél károsságának valószínűsége növekszik. Normál levelekben előforduló szavak egy vizsgált levélben a valószínűségi értéket csökkentik. Minél nagyobb a minta, minél több levelet vizsgált már át a szűrő annál pontosabb lehet a becslés. A Bayes analizátor egy valószínűségi értéket ad, az adminisztrátor feladata a káros/normális küszöbszint beállítása. Természetesen a spammerek is fejlődtek, a friss kéretlen levelek igyekeznek különféle módszerekkel kijátszani a spam szűrőket (például a levél szövegében normál tartalmat illesztenek, és egy grafikus csatolásban helyezik el a kéretlen reklámot).

Potenciális veszélyforrások felderítése, lokalizálása email vizsgálat és térinformatika segítségével

A vizsgálati módszer ismertetése előtt tekintsük át a módszert megalapozó tényeket:

- Napjainkban a kéretlen levelek túlnyomó többségéért a botnetek felelnek, vagyis egy kéretlen levél küldője 90% feletti valószínűséggel egy botnet tagja, vagyis egy fertőzött számítógép.
- A kéretlen levelet küldő számítógép TCP/IP protokollpárost használ, ami garantálja, hogy a feladó IP címe nem hamisított.
- A botnetek egyik célja a kéretlen levelek küldése. Emellett természetesen felhasználhatók DDoS támadásokra, felhasználó adatok eltulajdonítására, csalások elkövetésére is, de azt senki sem vitatja, hogy fontos az ellenük folytatott megelőző tevékenység (felderítés, azonosítás, fertőzöttség megszüntetése).
- Egy DDoS támadás elleni védekezés nehéz és költséges dolog. A védekezést megkönnyíti, ha közelítő adatokkal rendelkezünk arról, merről várhatók támadások, így az előzetes intézkedések korán megtehetők.

Az elektronikus levelezés nagy része a kéretlen kategóriába tartozik, ilyen módon rengeteg kárt okozva a továbbításban részt vevő infrastruktúra üzemeltetőinek, illetve a címzetteknek. Az infrastruktúrába az adatátviteli utak (amelyek hasznos sávszélessége emiatt lényegesen alacsonyabb lesz) mellett a levélszűrésre alkalmas számítógépek üzemeltetési költsége is beletartozik, nem elhanyagolható megtakarítást jelent tehát a botnetek felszámolása. Ennek első lépése a felderítés, amire a kéretlen levelek feladóinak IP cím vizsgálata is alkalmas, ha elfogadjuk azt a tényt, hogy a kéretlen levelet nagy valószínűséggel egy botnet tagja küldte. A módszer előnye az, hogy nem kell csali rendszereket (honeypot) telepíteni, elegendő az amúgy is rendelkezésre álló adatokból dolgozni, a keletkezett adatok alapján aztán lokalizálni is lehet a fertőzött gépeket. A Microsoft kutatói több botnet felderítési kísérletet is végeztek már a Hotmail levelezőrendszerben detektált kéretlen levelek alapján [7], [8]. Az általuk végzett vizsgálat főként a hálózati szintre korlátozódott, a rendelkezésükre álló hatalmas adatmennyiség segítségével pontos és többrétű elemzést tudtak lefolytatni. A módszer a levelek fejlécének vizsgálatán alapszik. A levél fejléce a címzett levelezőszerverére történő beérkezés után így néz ki:

```
Received: from 83.110.251.9 (dxb-b114995.alshamil.net.ae [83.110.251.9])  by mail.xy.hu  with ESMTP id m347L3XP014337;
          Fri, 4 Apr 2008 09:21:04 +0200
Message-ID: <000901c89624$051e9a4d$29abf7ab@fpfqfu>
From: "bourke kara" <*******@mail.ihs.gov>
To: <cimzett@xy.hu>
Subject:  Paris and Linsey lesbian video
Date: Fri, 04 Apr 2008 05:33:32 +0000
MIME-Version: 1.0
Content-Type: multipart/alternative;  boundary="----=_NextPart_000_0006_01C89624.051C0183"
X-Priority: 3
X-MSMail-Priority: Normal
X-Mailer: Microsoft Outlook Express 6.00.2900.3138
X-MimeOLE: Produced By Microsoft MimeOLE V6.00.2900.3198
```

Az első két sorban található adatokat a levelet fogadó szerver illesztette a fejléchez, ezekben található a feladó számítógép IP címével kapcsolatos összes adat. Látható, hogy a példában szereplő IP cím egy „.ae" végű domain névhez tartozó hálózat – amelyet az Egyesült Arab Emirátusokban regisztráltak – tagja, míg a feladó email címe az Egyesült Államokban működő IHS (Indian Health Service) szervezet tulajdona.
A fenti kéretlen levél fejléc vizsgálatával látszólag könnyen azonosítható a feladó fizikai elhelyezkedése, hiszen a szerver képes volt meghatározni a domain nevet. A felismerés azonban hibás eredményt is adhat:

- ha a küldő IP címéhez nem tartozik Reverse DNS bejegyzés, akkor a domain név nem meghatározható. Ez esetben más módszert kell találni a hálózat tulajdonosának és fizikai elhelyezkedésének meghatározására.
- a küldő IP címéhez tartozó szervezet meghatározása még nem feltétlenül utal a tényleges elhelyezkedésre. Elképzelhetők olyan hálózatok, amelyeket – bár ugyanaz a tulajdonosa - kisebb részekre darabolnak, és ezek különböző helyeken (akár különböző országokban is) üzemelnek.
- A domain név nem feltétlenül utal a végpont földrajzi elhelyezkedésére (egy .hu végű domain a világ bármelyik részén lehet).

Az ellenintézkedések - a kéretlen levelet küldő számítógép fertőzöttségének megszüntetése, vagy a hálózati forgalomból kiszűrése - végrehajtásához nagy segítséget nyújt, ha ismert a számítógép földrajzi elhelyezkedése. Idegen országban működő számítógép esetén lehetséges az adott ország megfelelő szervezeteit értesíteni, hazai hálózatok esetén pedig közvetlenül a hálózatot birtokló szolgáltatót. Sajnos az IP hálózatok címeinek kiosztása semmilyen kapcsolatban sincs az adott hálózat földrajzi elhelyezkedésével. Az IP címek kisebb, összefüggő területekből állnak, amely területeket bárki megvásárolhatja és használhatja. Ezek a területek különböző méretűek, ráadásul egy tulajdonos több, nem összefüggő területet is megvásárolhat. Szerencsére léteznek olyan adatbázisok, amelyek az IP címtartományokat és azok földrajzi elhelyezkedését tartalmazzák. Az esettanulmányban a MaxMind nevű cég GeoIP adatbázisának június elsejei kiadását használtam fel, annak is a GeoLite City változatát. Ez utóbbi ingyenes termék (a GeoIP előfizetéses) és 4 105 731 hálózat adatait tartalmazza. A hálózatokat az IP címtartomány kezdete és vége azonosítja. Mindegyik, ebben szereplő IP hálózatot egy földrajzi helyhez rendelték, míg a földrajzi helyekhez (245 831 darab) a következő adatok tartoznak:

- országkód;
- városkód;
- városnév;
- szélességi fok;
- hosszúsági fok.

Vagyis, az adatbázis segítségével lehetséges egy IP címet városhoz és térképen megjeleníthető ponthoz rendelni. Az adatbázist előállító cég szerint az ingyenes változat 99,5% pontosságú ország, míg 79% pontosságú város szinten, 25 mérföldes körben értelmezve. Ezek az adatok az USA hálózataira igazak, a világ többi részére nincs ilyen mérőszám. Természetesen léteznek más, hasonló célt szolgáló adatbázisok is [2], [3], sőt a magyar szolgáltatókra érvényes változatot is lehetséges lenne összeállítani.

Bár a térinformatikai rendszereket többnyire nem ilyen célra alkalmazzák, a jelenlegi, fejlett változatok egyszerű módszerekkel képesek többféle bemeneti formátumot használva együttműködni más rendszerekkel. A bemeneti formátumnak megfelelő állományok segítségével lehet pontokat kijelölni, amit az alkalmazás aztán elhelyez a térképen. Az együttműködéshez az is szükséges, hogy a Föld gömbölyű voltából adódó ábrázolási problémákat azonos módon kezeljék. A GeoIP a WGS 84 (World Geodetic System) ábrázolásmódot használja, tehát a megfelelő együttműködéshez olyan térinformatikai alkalmazás szükséges, amely ismeri ezt a szabványt. A komoly tudással rendelkező fizetős alkalmazások fő problémáját a felhasználható térképek választéka jelenti, de könnyen találhatunk ingyenes, egyszerű programozó felülettel rendelkező megoldást is. A teljes, részletes világtérképhez jutás problémáját egyszerűen kiküszöbölhetjük a weben elérhető térinformatikai megoldásokkal. A jelentősebb ilyen megoldások (a Microsoft Bing Maps [4], a Yahoo Map [5] és a Google Maps [6]) mindegyike rendelkezik programozói felülettel (API), amely segítségével bárki készíthet olyan térinformatikai alkalmazást, amely hozzáfér a teljes térképháttérhez, sőt, még műholdképekhez is. Hátrányuk, hogy csak online kapcsolat segítségével elérhetők, mivel a térképadatok az üzemeltetők számítógépes rendszereiben tárolódnak, a felhasználó csak a számára szükséges szeletekhez férhet hozzá.

Google Maps API

A feladat megoldása során választásom a Google Maps rendszerre esett, ennek oka főként a szélesebb támogatottság és emiatt a fellelhető példa alkalmazások nagyobb száma volt. Az API használata egyszerű: első lépésként egy kulcsot (egy szöveges azonosító, amely a felhasználási jogot adja meg a weboldal számára) kell igényelni, és máris használatba vehető a rendszer. Az alkalmazást JavaScript nyelven kell elkészíteni, a keretprogram a Google weboldaláról tölthető be, ez tartalmazza a térkép használatához szükséges osztályok (class) definícióit. A legfontosabb objektum a térkép objektum, amelyet a GMap2 osztályból lehet származtatni. Az objektumnak létrehozáskor meg kell adni azt a HTML szakasz azonosítót, amely a térképet fogja tartalmazni. A térképobjektumhoz tetszőleges számú vezérlőeszköz (nagyítás, kicsinyítés, térképtípus váltó gombok, stb…) adható, így a kezelőfelület is könnyen az igényekhez szabható.
A térkép definiálását és megjelenítését a következő függvény végzi:

```
function map_on() {
      if (GBrowserIsCompatible()) {
            map = new GMap2(document.getElementById("map"));
            map.setMapType(G_NORMAL_MAP);
            map.addControl(new GLargeMapControl());
            map.addControl(new GScaleControl());
            map.enableScrollWheelZoom();
            map.disableDoubleClickZoom();
      }
}
```

A térképen elhelyezkedő pontok megkereséséhez a legegyszerűbb módszer a szélességi és hosszúsági adatok megadása, ehhez a „GLatLng(szélesség, hosszúság)” osztály nyújt segítséget. A szélességi és hosszúsági értékek fokokban, tizedes tört alakban (tehát nem szögmásodperc, hanem tized, század, ezred fok) adhatók meg.
Az így keletkező földrajzi pont objektum aztán elhelyezhető a térképen. Ha ezt a földrajzi pontot meg szeretnénk jelölni a térképen, akkor egy markert kell elhelyezni rajta, amihez rendelhetünk grafikát, a földrajzi koordinátáit, illetve egy eseménykezelőt, ami a különböző felhasználói aktivitást szolgálják ki (kattintás, dupla kattintás, vonszolás). A spam küldő bot

kliensek koordinátáit ilyen markerek segítségével helyeztem el a térképen, a következő függvény segítségével:

```
function insertPoint(latitude, longitude, comment) {
      var coord = new GLatLng(latitude,longitude);
      map.setCenter(coord);
      var baseIcon = new GIcon();
      baseIcon.iconSize=new GSize(32,32);
      baseIcon.shadowSize=new GSize(56,32);
      baseIcon.iconAnchor=new GPoint(16,32);
      baseIcon.infoWindowAnchor=new GPoint(16,0);
      var icon = new GIcon(baseIcon, "marker.png", null, "shadow.png");
      var botloc = new GMarker(coord,icon);
      GEvent.addListener(botloc, "click", function() {
            marker.openInfoWindowHtml(comment);});
      map.addOverlay(botloc);
}
```

A markerek segítségével jelöltem be a térképen a botnet kliensek helyét, természetesen nem egyenként, hanem csoportosítva. A csoportosítás háromféle volt:

- Világméretű: a küldő számítógépek IP címeit a GeoIP adatbázisa alapján országonként csoportosította az alkalmazás, majd az országokban előforduló egyik város koordinátáját felhasználva helyezte el a markert. Lehetséges lett volna egyedi, országokra jellemző pontokat választani markerhelynek, de ez szükségtelenül sok plusz munkát jelentett volna (a világ összes országában be kellett volna jelölni a fővárost).
- Európai szintű: a küldő számítógépek IP címei közül az alkalmazás kigyűjtötte az európai országokat, majd országonként összesítette. A marker elhelyezése az adott ország fővárosának koordinátájára történt. Ehhez készítettem egy európai fővárosok adatait tartalmazó adatbázist.
- Magyarországi szintű: az alkalmazás csak a magyarországi hálózatokból érkezett kéretlen levelek IP címeit szűrte, majd városonként összegezte. A marker elhelyezése a városok koordinátája alapján történt.

Valamennyi mód esetében, áttekinthetőségi okokból csak a 100 legaktívabb küldő ábrázolása történt meg.

Esettanulmány

Az elmélet igazolására végeztem egy kísérletet, saját adatokra támaszkodva. Saját levelezőszerveremen a spamszűrő több éve végzi feladatát, így rendelkezésemre állt több milliónyi elektronikus levél szűrési naplója. A feldolgozás első részében egy erre a célra írt egyszerű script segítségével a tárolt elektronikus levél fejlécek alapján egy SQL adatbázis táblába gyűjtöttem a levelek feladóinak (akiket így potenciálisan botnet tagnak tekintettem) IP címeit. Takarékossági okokból csak a 2007. január és 2009. május közti adatokat dolgoztam fel, a rekordok száma még így is 1184620 lett. A címek közül töröltem az ismertebb ingyenes levelezőrendszerekből érkezetteket. Ennek oka az, hogy 2008-ban spammerek sikeresen áttörték az automatikus, gépek által végzett regisztrációt megakadályozó CAPTCHA (a Completely Automated Public Turing Test To Tell Computers and Humans Apart kifejezés rövidítése) védelmi rendszert [10]. A GeoIP hálózati tartományokat tartalmazó adatbázisának segítségével lehetségessé vált az IP címekhez földrajzi elhelyezkedést rendelni. A hozzárendelés két lépcsős, a hálózatok adatbázisa csak egy földrajzi hely azonosítót tartalmaz, ez az azonosító aztán egy másik adatbázisban (location) a tényleges földrajzi adatokra mutat. Itt merült fel az első probléma, a rengeteg adat és a nehezen optimalizálható keresési metódus miatt a második lépés feldolgozási időigénye hatalmasra duzzadt, így valós idejű statisztikák

készítésére csak elfogadhatatlanul nagy erőforrás felhasználás mellett lett volna képes. Ezért ezt a lépést nem a statisztika készítésekor végzi az alkalmazás, hanem az IP címek feldolgozása során a rekordok egészülnek ki egy location ID mezővel, amelynek feltöltése is ekkor történik meg.
Az így előálló adatbázis már könnyen és viszonylag gyorsan kezelhető, egy egyszerű PHP script képes a kívánt adatokat kigyűjteni, majd a megjelenítést végző JavaScript programnak átadni. A próbarendszer felületét az alábbi ábra mutatja:

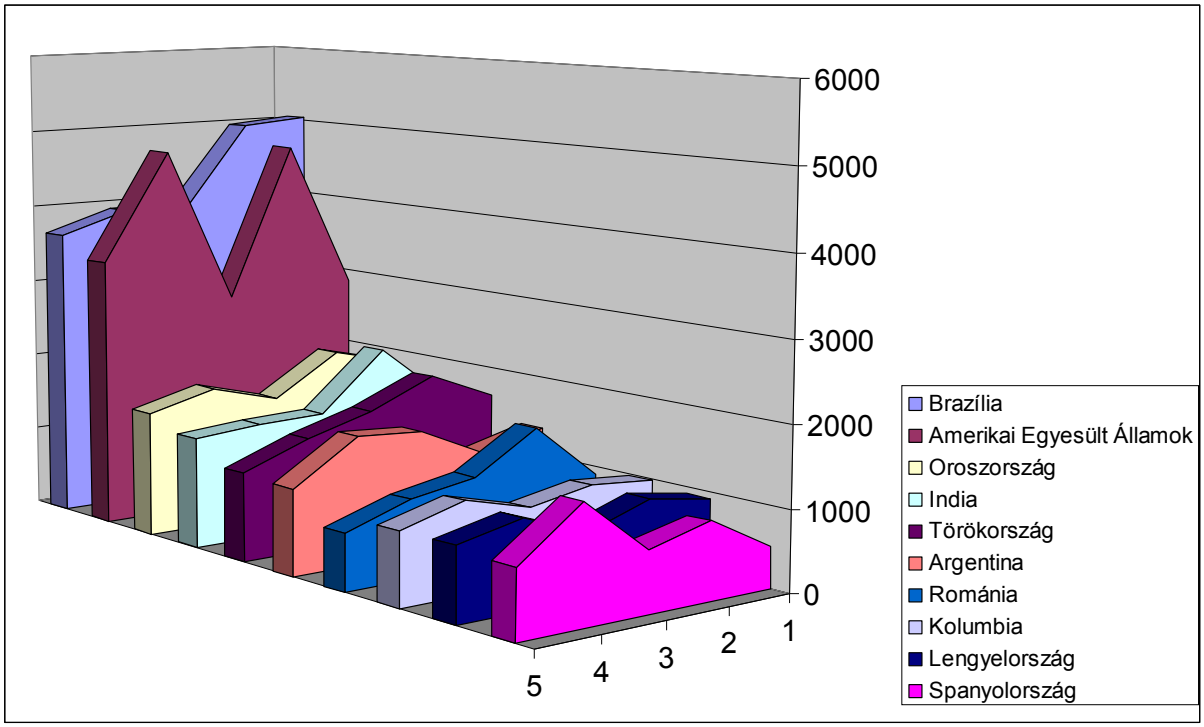
Terület: Magyarország Év: 2009 Hónap: Május Frissítés

Város	Szélesség	Hosszúság	Darab
Budapest	47.500000	19.083300	139
Oroszlány	47.483300	18.316700	15
Szeged	46.250000	20.166700	7
Ács	47.700000	18.016700	6
Rakamaz	48.133300	21.466700	5
Kazincbarcika	48.250000	20.633300	4
Jászberény	47.500000	19.916700	3
Dunakeszi	47.633300	19.133300	3
Budaörs	47.450000	18.966700	3
Vecsés	47.400000	19.283300	2
Székesfehérvár	47.200000	18.416700	2
Esztergom	47.800000	18.750000	2
Kiskunfélegyháza	46.713600	19.852200	2
Edelény	48.300000	20.733300	2

Fertőzöttségi térképek

Az elkészült alkalmazás segítségével gyorsan és egyszerűen lehet fertőzöttségi térképeket készíteni, amelyek a fertőzött végpontok földrajzi elhelyezkedését mutatják:

Emellett a GeoIP, és a folyamatosan feldolgozott IP címek segítségével országonkénti megoszlást is lehet számítani az eltelt idő függvényében, amely segítségével a tendenciák nyomon követhetők.



Meglepő módon a statisztikát Brazília vezeti, az Egyesült Államok és Oroszország előtt. A fenti ábra a 2009. évből eltelt 5 hónap adatai, 158000 darab IP cím feldolgozása alapján készült, természetesen nem általános érvényű.

Összegzés

A fentiekben igyekeztem rövid összefoglalást nyújtani a botnetek elleni küzdelem egyik módszeréről, amely segítségével azonosíthatók a fertőzött végpontok. Az azonosítás a levelezőszerverek kéretlen levelek szűrése után keletkező naplóállományaiból elvégezhető, így járulékos befektetést, újabb eszközök beüzemelését nem igényli. A kiszűrt végpontok földrajzi helyének megkereséséhez a Maxmind cég GeoIP adatbázisát használtam, amely egy IP adatbázis, az egyes hálózati tartományokhoz tartozó földrajzi pozíció adatokkal.

A botnetek egyre nagyobb veszélyt jelentenek, ezért az ellenük folytatott harcban sokkal nagyobb figyelmet kell szentelni a megelőzésnek, aminek egyik eszköze lehet a térinformatika. A reaktív, botnetek elleni rendszereknek természetesen nem szükséges tényleges földrajzi koordinátákkal dolgozniuk (egy megfelelő pontosságú adatbázis karbantartása szinte lehetetlen feladat lenne), elegendő az adott IP címet tartalmazó hálózat tulajdonosának paramétereit szolgáltatniuk. Ilyen adatbázisok megfelelő pontosságban rendelkezésre állnak, az ismertetett rendszer kis módosítással ezeket az adatokat is képes feldolgozni. A hálózat tulajdonosának ismeretében aztán kellő gyorsasággal semlegesíteni lehet a potenciális veszélyforrást jelentő fertőzött számítógépet.

Felhasznált irodalom

[1] http://www.maxmind.com

[2] http://www.hostip.info/dl/index.html

[3] http://www.ipligence.com/products

[4] http://msdn.microsoft.com/en-us/library/dd877180.aspx

[5] http://developer.yahoo.com/maps/

[6] http://code.google.com/intl/hu-HU/apis/maps/

[7] S-GPS: Spammer Global Positioning System - Microsoft Research
http://research.microsoft.com/en-us/projects/S-GPS/

[8] Yao Zhaoy, Yinglian Xie, Fang Yu, Qifa Ke, Yuan Yu, Yan Cheny, and Eliot Gillumz: BotGraph: Large Scale Spamming Botnet Detection
http://www.usenix.org/event/leet08/tech/full_papers/zhuang/zhuang.pdf

[9] PTA CERT-Hungary Központ éves jelentés

[10] Google's CAPTCHA busted in recent spammer tactics
http://securitylabs.websense.com/content/Blogs/2919.aspx

Valamennyi internetes forrás 2009. június 14-én elérhető volt.