Biztonsági sérülékenységek javításának empirikus vizsgálata nyílt forrású rendszerekben

Mosolygó Balázs József
III. évf programtervező informatikus BSc

Vándor Norbert Rudolf
III. évf programtervező informatikus BSc

Témavezetők: Dr. Hegedűs Péter, Antal Gábor

SZTE TTIK Szoftverfejlesztés Tanszék

Napjainkban az informatikai rendszerek egyre szélesebb körben való elterjedése miatt a folyamatosan növekvő felhasználói bázis még sosem tapasztalt nyomást gyakorol a programok készítőire. Annak érdekében, hogy a fejlesztők lépést tudjanak tartani a termékeik iránt egyre növekvő igényekkel, kénytelenek már kész megoldásokhoz folyamodni. A nyílt forráskód bázisok rendkívüli lehetőséget biztosítanak ehhez, mindazonáltal komoly veszélyeket is rejthetnek. A széles körben adaptált kódbázisok fejlesztőire még nagyobb felelősség hárul, hiszen ha hibáznak, az akár a teljes közösségre is hatással lehet. Ezért kitüntetett jelentősége van a nyílt forrású rendszerekben jelen lévő biztonsági hibáknak, amelyek hatalmas károkat okozhatnak azon szoftverrendszerek esetén, melyek támaszkodnak ilyen nyílt forrású programokra. Kutatási munkánk során az egyes nyelveken írt nyílt forrású programokban tipikus biztonsági hibákat vizsgáljuk annak érdekében, hogy hatékonyabbá tehessük az azok ellen folytatott küzdelmet. Ehhez repository bányászó technikát alkalmaztunk, és a GitHub-on elérhető projektek commit adatait elemezve létrehoztunk egy adatbázist, ami népszerű nyelvek nagyobb projektjeinek fejlesztési adatait tartalmazza. Ezen adatok elemzése által betekintést nyerhetünk a közösségek általános fejlesztési mechanizmusaiba. Azt találtuk, hogy a több aktív fejlesztővel rendelkező projektek nem feltétlenül hatékonyabbak a hibák javításával kapcsolatban. Továbbá, nyelvtől és súlyosságától függetlenül a közösségek általában ugyanannyi idő alatt reagálnak a különböző sérülékenységtípusok. Valamint, hogy bizonyos nyelvekkel dolgozó közösségek gyakrabban követnek el bizonyos típusú hibákat mint mások, azaz vannak jellemző hibáik. Az adatbázis által tárolt információk további célokra is felhasználhatóak, munkánk leginkább csak demonstrálja az összegyűjtött adatokban rejlő potenciált. Olyan további felhasználási lehetőségek tárházát kínálják, mint például hiba előrejelző modellek tanítása.