Интелигентни методи и киберсигурност

Румен ТРИФОНОВ*, Огнян НАКОВ*, Пламен ВАЧКОВ, Славчо МАНОЛОВ*, Радослав ЙОШИНОВ**, Георги ПОПОВ*, Георги ЦОЧЕВ* и Галя ПАВЛОВА***

* Технически Университет - София, Факултет по Компютърни системи и технологии, София 1000, България, бул. „Кл. Охридски“ 8, бл. 1, e-mail: r_trifonov@tu-sofia.bg

** Българска академия на науките, Лаборатория по телематика, София 1000, България, ул. „Акад. Георги Бончев”, 8, e-mail: yoshinov@cc.bas.bg

Резюме. *В условията на пето поколение киберпрестъпност, характеризиращо се с автоматизация на разработването и разпространението на инструментите за атака, както и интеграция в рамките на няколко комплекта инструменти, преобладаващото количество експерти считат, че традиционните методи за защита са вече слабо продуктивни и е необходим качествен преход към нови инструменти за реализация на мрежовата и информационна сигурност. Едно от посочваните с приоритет направления на този преход е широкото приложение на интелигентни методи за анализ на обменяната на информация, на потоците в мрежите, на източниците на заплахи, както и планиране на ефективни мерки за въздействие, в т.ч. проактивни. Настоящият доклад е посветен на приложението и експериментирането на един конкретен метод на изкуствения интелект за защита на мрежови сървър и хостове в мрежата, изпълнявани в катедра „Информационни технологии в индустрията“ на Факултета по компютърни системи и технологии на Техническия университет – София.*

In conditions of the fifth generation of cybercrime, characterized by the automation of the development and dissemination of attack instrument, as well as integration within several sets of instruments, the most of experts believe that the traditional methods of protection are not enough effective and that a qualitative transition to new instruments for implementation of network and information security is required. One of the priority directions of this transition is the widespread application of intelligent methods for analyzing the information exchange, the network flows, the sources of threats and planning effective impact measures, including proactive ones. This report is dedicated to the application and experimentation of a specific artificial intelligence method for protection of network servers and hosts in the network at the Department of Information Technologies in Industry, Faculty of Computer Systems and Technologies at the Technical University of Sofia.

Съвременни заплахи на киберпрестъпността

Съвременната икономика и общество се развиват все по-перспективно в нови направления, свързани с интензивното използване на информационните и телекомуникационни технологии, софтуерни системи за управление, както и на ефективни процеси, базирани на дигиталните инфраструктури. Към традиционните рискове се добавят нови, кибер-рискове с ключово значение, игнорирането на които може да доведе до катастрофални резултати.

Кибер-атаките са директна заплаха за сигурността на гражданите и функционирането на държавата, икономиката, обществото, науката и образованието. Те могат да бъдат извършени от разстояние, с прости и ефективни механизми, минимални икономически ресурси и да причинят значителни поражения с нанасяне на материални и дори човешки загуби. Кибер-атаките нямат национални, културни или юридически граници. Рисковете и заплахите в кибер-пространството са трудни за дефиниране поради сложността за определяне на източника на въздействие, целите и мотивите, сложността и интензивността на съвременните комуникационни и информационни процеси, динамиката на логическите и физическите връзки и неопределеността на

процесите. Сред най-сериозните деструктивни въздействия са тези от хибриден характер - комбинация от кибер-атака и физическа атака, кибер-атака целяща критичен кинетичен процес, кибер-атака по време на природно бедствие или неизправност в критични системи.

Атакуващите сега консолидират активите си с цел създаване на глобални мрежи, които поддържат координирани престъпни действия.

Терминът „кибер-престъпност" обикновено се ограничава до описване на престъпна дейност, в която компютър или мрежа се явяват съществена част от престъпността [1]. В по-разширен аспект този термин включва традиционни престъпления, в които компютри или мрежи се използват, за да се даде възможност на извършване на незаконни дейности. При това, компютърът или мрежата могат да бъдат както инструмент за престъплението, така и обект на престъплението.

На световната икономика й бяха необходими няколко века от утвърждаването на пазарните отношения до постигане на глобализация. За сенчестата икономика на кибер-престъпността за този преход бяха достатъчни едно-две десетилетия. Тийнеджърите, пишещи вирусни програми главно за самоутвърждаване, бързо отстъпиха място на добре организирани картели. По оценки на сериозни анализатори годишният оборот на „Интернет-базираната сенчеста икономика" вече се приближава до търговията с наркотици. Налице е сложен „он-лайн" черен пазар с десетки хиляди участници и без национални граници.

Кибер-заплахите са асиметрични, нелегални и силно наподобяват похватите на класическия тероризъм. Т.е. отделно лице или малка група някъде по света могат без особени разходи да се опитат да проникнат в системи, съдържащи жизнено важна информация или да предприемат разрушаващи атаки срещу критична инфраструктура. Инструментите и ресурсите за такива атаки са лесно достъпни чрез Интернет и уязвимостите на атакуваните системи са все по-лесно откриваеми и използваеми.

Освен това кибер-престъпниците използват за своите цели множество нищо не подозиращи потребители на компютри и Интернет. Заразените компютри на тези потребители формират т.н. „бот-мрежи", чрез които се организират атаки към важни източници на информация или се изпращат „спам-писма" и вредни програмни кодове към многобройни адресанти, а също така се разпространява забранено съдържание, например, детска порнография [1, 2].

Първото поколение на престъпните действия в кибер-пространството се характеризира с бързото размножаване на червеи, които експлоатират разпространените уязвимости. Заплахи с голяма степен на въздействие от това поколение бяха червеи, които в своята съвкупност повредиха милиони компютри по целия свят. За киберпрестъпниците от първото поколение приоритет № 1 е да ги забележат.

Отличителната черта на кибер-престъпленията от второ поколение е мотивът за печалба. Осъзнаването, че хакерство може лесно да бъде използвано за парична печалба, накара редиците на хакерите да набъбнат с хора, които искат да спечелят. Ботнетите (големи мрежи от заразени компютри) се превърнаха в предпочитано оръжие за кибер-престъпниците, позволявайки им да „изпомпват" милиони спам-съобщения или да извършат атаки от типа „разпределен отказ на услуги (DDoS - Distributed Denial of Service)" върху бизнеса или администрациите.

Две отличителни черти отбелязва третото поколение кибер-престъпления: организация и дискретност. Кибер-престъпниците стават по-зрели, като осъзнават предимствата на съвместната работа за незаконни доходи. Освен това те си поставят по-високо печеливши цели. Методите (червеи, троянски коне, DDoS, ботнети и т.н.) са същите като в предишното поколение, но изпълнението отразява влиянието на традиционните за криминалния свят начинания.

Това поколение е насочено преди всичко към фирмите, които боравят с големи суми пари, като финансови институции и организатори на хазартни игри.

Възникването на дейност „C2C (Criminal-to-Criminal)" дава началото на четвъртото поколение на кибер-престъпността. Появява се силна и ефективна сива икономика, която предоставя възможности за киберпрестъпниците да купуват и продават стоки и услуги един на друг. Отделни специализирани киберпрестъпни бизнеси станаха известни, включително:

- „аукционни къщи" - електронни пазарища, където киберпрестъпниците купуват и продават изпълним код, включително зловредни програми за софтуерни уязвимости, които не са публично известни;
- услуги за разпространение на зловреден софтуер, който да зарази хиляди хостове. Тези услуги обикновено имат установена среда за

разпространение, например, мрежа от заразени уеб сайтове или заразени онлайн реклами;

- ботнети под наем, които поддържат един или повече ботнет, наемани от други киберпрестъпници. Наетия ботнет може да се използва за изпращане на спам, за хостване на нелегитимни сайтове, за кражба на чувствителна информация, за изпълнение на DDoS атаки и провеждане на много други престъпни дейности;
- продавачи на самоличност от ново поколение, които организират покупки и продажби на откраднати данни. Тези нови услуги дават на киберпрестъпниците „он-лайн" платформа за покупка, продажба и управление на портфейл от откраднати записи – това е нещо като нелегална „он-лайн" борса, която помага на хакерите да максимизират своите „инвестиции";
- лицензиран зловреден код - авторите на злонамерен код приемат лицензионни модели, принуждавайки по този начин други киберпрестъпници да плащат за техния зловреден софтуер. Това дава повече финансови възможности за авторите да усъвършенстват своите разработки, също така дава възможност на другите киберпрестъпници да закупуват зловреден софтуер от висока класа, вместо да се налага да се развиват свой;
- социални мрежи за киберпрестъпниците, които предоставят репутационна класация на купувачи, продавачи и партньори в престъпленията в кибернетичното пространство. Те включват "доверени" лица, които извършват т.н. „escrow" функции, когато един или повече "ненадежден" партньори участват в операции на престъпления в кибер-пространството.

Заплахите в петото поколение са все по-автоматизирани, като започват да се ползват от предимството на инструменти и техники за писане на скриптове за автоматизиране на различни етапи на техните схеми. По-малко опитни хакери могат да закупят инструменти за лесно идентифициране на уязвими цели, за компрометиране на системи и за кражба на данни. В някои случаи, в по-големите схеми за престъпления в кибер-пространството се наблюдава интеграция в рамките на няколко комплекта инструменти, които изпълняват различни функции.

Една от особеностите на петото поколение са т.н. "Модерни настойчиви заплахи (APT - Advanced Persistent Threats)", които станаха известни през 2010 г. като наименование за целенасочени атаки срещу конкретни организации от определени, добре координирани кибер-престъпници. В общността на експертите, APT най-често се отнася до сложни атаки, насочени към правителства и корпорации, с цел да се събира разузнавателна информация или постигане на конкретни финансови цели. APT често се използва от държавни органи или техни агенти. В някои случаи, те са свързани с терористични и сепаратистки политически групи.

Класификация на видовете атаки

В общия случай, заплахите могат да се групират условно на три надграждащи се нива [3]:

- „известни известни" – известни слабости, заплахи и пробиви, свързани с основната „триада" на информационната сигурност (КИД – конфиденциалност, интегритет, достъпност);
- „известни неизвестни" - комбинирани заплахи, свързани с информационната сигурност, разнообразие от APT, атаки срещу репутацията на организации и личности, кампании за дезинформация, пробиви в КИД в особено големи мащаби (национални, регионални и световни), изискващи разширено и системно прилагане на КИД за всички активи в дигиталната екосистемам
- „неизвестни неизвестни" – непредсказуеми, неочаквани заплахи в киберпространството, динамично променящи се рискове и комплексни въздействия с непредсказуеми последствия, които изискват гъвкавост и устойчивост на системите, организацията и процесите, и съответни изисквания при разработването и внедряването им - основните характеристики на състоянието кибер-устойчивост.

Инструменти и устройства за активно противодействие на заплахите срещу информационната сигурност

Идентифицирането на атаките е процес на откриване на проникващи събития, възникващи в процеса на експлоатация на дадена информационна система. Наличието на системи за идентифициране на атаки е задължителен елемент от политиките за сигурност. Аналогично на системите за управление на високоотговорни технологични процеси, в системите за защита на информационните системи възниква изискването за разпознаването на проникващите действия в момента на тяхното възникване, а не след тяхното реализиране. Едновременно с откриването на опит за проникване е необходимо да започне функционирането на механизъм за превантивни действия, които са свързани с ограничаване или изолиране на действието на източник на атака и

предприемане на активно противодействие с цел неговото блокиране и привеждане в неработоспособно състояние [4].

Понастоящем като основни стълбове на рамката за сигурност на повечето организации функционират защитни стени, системи за предотвратяване на прониквания, антивирусен софтуер, уеб шлюзове и мрежови примамки [5].

Защитните стени и екраниращите рутери са средства за ограничаване на трафика между две мрежи или към и от даден хост. Целта на ограничаването е недопускането на определен трафик да бъде предаден от дадено мрежово устройство или да бъде получен от него. Ограничаването на трафика се извършва чрез предварително зададени дефиниции, които се прилагат върху пакетите, принадлежащи на съответния източник на трафик. Всяко правило определя дали пакетът може да постъпи или напусне съответната крайна система (хост или мрежово устройство).

Защитната стена е устройство, използвано като механизъм за контрол на достъпа до конкретна мрежа или набор от мрежи. В повечето случаи защитните стени служат да предотвратят достъпа на външни лица до локалната мрежа. Но те могат да служат и за създаване на по-сигурни участъци вътре във локалната мрежа за особено важни функции. Наличието на някакъв тип защитна стена е задължително, ако мрежата разполага с връзка към Интернет, независимо дали тази връзка е широколентова (кабелен модем или DSL), T1 или някакъв друг вид високоскоростна връзка.

Екраниращият рутер сканира и анализира стойността на полетата от заглавните части на пакетите. Рутерът се използван за да отдели локалната мрежа от Интернет. При това разделяне се дефинират две зони: зона на риск; зона на сигурност, доверена зона.

Интернет шлюзът обикновено, е софтуер, който организира достъпа до Интернет. Той е работен инструмент за системните администратори, позволяващ да се контролира трафика и действията на служителите, като разпределя достъпа до потребителите, следи трафика, ограничава достъпа на отделни потребители или групи потребители до ресурсите на системата. Интернет шлюзът може да съдържа прокси сървър**, защитна стена**, антивирусни и други услуги на мрежата [6].

Антивирусният софтуер се използва за предотвратяване, разкриване и премахване на зловреден софтуер, включително, компютърни вируси, червеи, троянски коне, шпионски и рекламен софтуер.

Системата за откриване на проникване (IDS - Intrusion Detection System) е софтуер или устройство, което автоматизира процеса на откриване на проникване. Система за предотвратяване на прониквания (IPS - Intrusion Prevention System) е софтуер или устройство, което има всички възможности на IDS и също може да се опита да спре евентуални инциденти. IDS и IPS технологиите предлагат много съвпадащи възможности и администраторите могат в дадени случаи да блокират превантивните функции в продуктите на IPS, карайки ги да функционират като IDS. При това положение, може за краткост да се използва термина (IDPS - Intrusion Detection and Prevention Systems) [7].

Много IDPS могат също така да идентифицират разузнавателна дейност, която да посочи, че атаката е предстояща. Например, някои инструменти и форми на зловредния софтуер, в частност червеи, изпълняват разузнавателни дейности като сканират хостовете, за да определят цели за следващите атаки. IDPS може да бъде в състояние да блокира разузнаването и да уведомява администраторите за сигурност, които могат да предприемат действия, ако е необходимо, за да се предизвикат други проверки за сигурност с цел предотвратяване на инциденти. Тъй като разузнавателната дейност е доста често явление в Интернет, откриването й често се извършва предимно върху защитени вътрешни мрежи.

Мрежовата примамка (Honeypot) е система за откриване на опити за неправомерен достъп до информационни ресурси. Honeypot имитира работата на реалната система, която се явява потенциална цел на атаката и на неправомерния достъп. Тя отвлича върху себе си вниманието и ресурсите на нарушителя, фиксира всички негови действия и информира органите по информационна сигурност за фактите на нарушенията. При това, в зависимост от типа на мрежовата примамка, могат да се имитират всякакви ресурси, явяващи се потенциални обекти за атаки: сървъри, бази данни, мрежови услуги, файлови ресурси и др.

Предимствата на Honeypot-системите се определят от самия принцип на тяхната работа. Преди всичко, това е практически пълното отсъствие на лъжливи сработвания. Тъй като мрежовата примамка само имитира реалната система и към нея не се обръщат нито реалните

ползватели на мрежата, нито реалните мрежови приложения, то всякаква активност на Honeypot и всеки опит за обръщане към него е незаконен и свидетелства или за атака или за изследване на мрежата от нарушител с цел да се намерят уязвими места в нейната защита.

Необходимост от въвеждане на допълнителна интелигентност на защитните устройства

Въпреки значимостта на посочените защитни средства и антивирусни програми, тяхната комбинация в съвременните условия на комбинирани атаки често не е достатъчна, за да се предотвратят атаките на съвременните версии на зловреден софтуер и опити за неправомерно проникване. Новото поколение защитни стени не са достатъчно ефикасни, тъй като те са оптимизирани за прилагане на политики, свързани с традиционните технологии, а не за откриване и блокиране на бързо променящи се заплахи.Ето защо експертите смятат, че е необходим качествен преход към нови инструменти за реализация на мрежовата и информационна сигурност.

Едно от посочваните с приоритет направления на този преход е широкото приложение на интелигентни методи за анализ на обменяната информация, на потоците в мрежите, на източниците на заплахи, както и планиране на ефективни мерки за въздействие, в т.ч. проактивни (т.е. атакуващи основните източници на заплаха, такива като управляващи центрове на „бот-нет“-и пр.).

Световната практика отбелязва вече значителен брой от разнообразни приложения на „изкуствен интелект“ в компютърната сигурност. Без да се прави опит за изчерпателна класификация, бихме могли да разделим тези приложения в две основни направления [3, 4]:

А. Условно наречени „разпределени“ или „мрежови“ методи:

А1. Мулти-агентни системи от интелигентни агенти;

А2. Невронни мрежи;

А3. Изкуствени имунни системи и генетични алгоритми и т.н;

Б. Условно наречени „компактни“ методи:

Б1. Системи за машинно самообучение (Machine Learning), в т.ч.: асоциативни методи, индуктивно логическо програмиране, Бейсова класификация и пр.

Б2. Алгоритми за разпознаване на образи;

Б3. Експертни системи;

Б4. Размита логика и пр.

Наличните академични ресурси показват многобройни реализирани приложения в борбата с компютърните престъпления. Невронните мрежи основно са приложени за откриване и предотвратяване на опити за проникване и профилактика, но има и предложения за използване на невронни мрежи в борбата с "DoS“ (Denial of Service - отказ от обслужване)-атаки, за откриване на вируси от типа червей, за откриване на спам, за класификация на злонамерен софтуер, както и за съдебни (Forensics) разследвания. Например, т.н. NeuroNet е показала експериментално добра ефективност срещу разпределени атаки „отказ от обслужване“, насочени срещу TCP-нивото на мрежата.

Мулти-агентните интелигентни системи са използвани основно в системите за откриване и прадотвратяване на проникване (IDS/IPS), както и за адаптивно настройване на защитни стени. Мобилни интелигентни аганти могат да участват в трафика между основни транспортни пунктове в мрежата, за да открият подозрителна кибер дейност. Пример за такова приложение е SAaaS (Security Audit as a Service) – система за откриване на инциденти в облачни приложения, базирана на автономни интелигентни агенти.

Изкуствените имунни системи, аналогично на биологичните, са предназначени да поддържат стабилност в променяща се среда. Имунно-базираното откриване и предотвратяване на проникване включва еволюция на имуноцитите и създаване на антигени за откриване и премахване. Интелигентните имунни системи се използват основно в новото поколение анти-вирусни алгоритми, както и за екстракция на спам. Като пример може да бъде посочена SGDIDS (Smart Grid Distributed Intrusion Detection System) – йерархична система, откриваща и класифицираща зловредни кодове и възможни кибер- атаки. Резултатите от симулацията й са показали нейната приложимост за подобряване сигурността на мрежата.

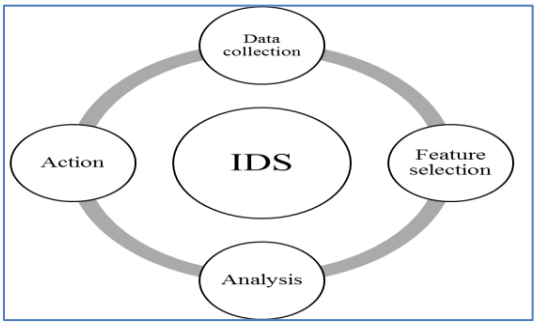
Също като пример може да бъде посочен т.н. Хомеостатичен модел за координация на сигурността на система за е-Правителство, базиран на невро-ендокринно-имунен принцип. Изследването на прототипа му показва, че изкуствен модел на хомеостаза може да интегрира различни продукти за сигурност, като ги координира в откриване на проникване и предотвратяване на потенциални атаки или експлоатиране на уязвимости.

Заслужава да се спомене новия метод за откриване на проникване, базиран на т.н. „размита мрежа (Fuzzy Network)", който използва асоциативно издирване в генетично програмирана мрежа. Предложеният метод се е оказал гъвкав и ефективен при откриване на аномалии в мрежи и способен да работи със смесени бази данни, които съдържат дискретни и непрекъснати атрибути [8].

Реален пример на разработка и експериментиране на приложение на изкуствения интелект в информационната сигурност

Понастоящем, в катедра „Информационни технологии в индустрията", Факултет по компютърни системи и технологии на ТУ – София се изпълнява проект „Повишаване нивото на мрежовата и информационна сигурност чрез използване на интелигентни методи", финансиран от фонд „Научни изследвания" на Министерството на образованието и науката. Извършено е моделиране и експериментално изследване на избрани след изчерпателен анализ определени интелигентни методи, надграждащи функционалността на системи за активно противодействие на опити за проникване при защита на мрежови сървър и хостове в мрежата.

Най-популярната в последно време система за активно противодействие е системата за откриване и противодействие на проникване (IDPS – Intrusion Detection and Prevention System). IDPS се състоят от четири основни елемента (функции): за събиране на данни, за селекция, анализ и действие (Фиг. 1) [9].



Фиг. 1. Система за откриване и противодействие на проникване

Събраните данни обикновено се записват в база данни, от която се изтеглят и анализират. Анализът, като правило, се извършва чрез сравняване на данните с определени модели или сигнатури. Друг метод се основава на откриване на аномалии в работата на средата. Действието определя атаката и реакцията на системата.

Управлението на IDPS се осъществява, чрез централизирано устройство, което получава информация от сензорите, изпълнява анализ на информацията за събития, които сензорите са предоставили и може да идентифицира определени събития. Съвпадението на информация за събития от множество сензори, като например намирането на събития, предизвикани от един и същ IP адрес, е известно като корелация. Управлението може да се осъществява от софтуерен продукт и/или хардуерна машина. Сървърът с базата данни е хранилище с информация за отделните събития, записани от сензорите или устройствата за управление.

Изборът на приложение на изкуствения интелект, което да бъде включено в експеримента, е направен въз основа на анализ на източниците на информация относно подобни приложения и на резултатите от подобни експеименти.

Традиционно, подходите за откриване и превенция на проникване се базират на два основни принципа – откриване на аномалия и откриване злоупотреба, макар че между тях не съществува значителна разлика в характеристиките. Механизмът за откриване на опити за проникване обикновено разкрива следните подозрителни действия в мрежата: а) опити да се използват услуги, блокирани от защитни стени; б) неочаквани заявки, особено от непознати адреси; в) неочаквани шифровани съобщения; г) твърде активен трафик от непознати сървъри и устройства; д) значителни промени от предишните действия на мрежата; е) опити за използване на известни бъгове или уязвимости; ж) опити за достъп от непознати потребители от неочаквани адреси; з) неправилно или подозрително използване на администраторски функции; и) значителни изменения в обичайните действия на потребителя и т.н.

Видът на откриването на атаката зависи от характера на използваните заплахи (познати, непознати и комбинация от двата вида). Създадени са критерии за оценка на ефективността на откриването и нивото на изпълнение на противодействието. От изключително важно значение е и постигането на правилния баланс между неверни положителни резултати и фалшиви негативни. Неверните положителни резултати (т.н. фалшиви аларми) може да се окажат не по-малко вредни, отколкото грешните отрицателни резултати.

Направен е анализ-сравнение на различни методи на изкуствения интелект от гледна точка на горепосочените критерии. Установено е, че методологията на аномално настроените мулти-

агентни системи превъзхожда болшинството традиционни системи, базирани на изкуствения интелект при откриването на атаки, особено с неизвестен характер. Ефективността на откриване на опасности при мулти- агентните системи също превъзхожда традиционните системи. Най-съществените аспекти на мулти-агентно-базираните системи за IDPS са високата точност, самообучението и устойчивостта.

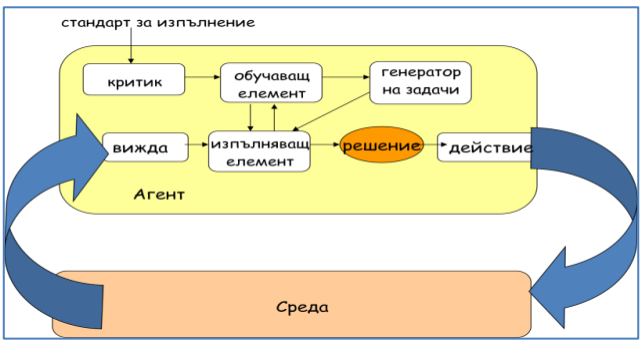
Описаните в източниците практики показват [10, 11], че резултатите относно правилното откриване на заплахи с използването на мулти-агентно-базирани системи, постоянно се увеличава, като процентът на лъжливите аларми драстично намалява. Без съмнение, мулти-агентно-базираните подходи могат потенциално да достигнат повишена гъвкавост, което ще ги направи още по-популярни в близко бъдеще. Поради това, експерименталният модел, създаден на първия етап от проекта, е комбинация от мулти-агентна система и система за откриване и предотвратяване на проникване (IDPS).

Автономните агенти са изчислителни системи, които съществуват в сложна динамична среда, действат самостоятелно в тази среда и по този начин реализират набор от цели и задачи, за които те са проектирани. Агентите притежават следните основни характеристики:

- автономност - възможност да действа самостоятелно без пряката намеса на хора или други агенти и да има контрол върху собствените си действия и вътрешни състояния;
- социалност – възможност за взаимодействат помежду си като използват език за комуникация;
- реактивност – възприемане на заобикалящата ги среда и реагиране своевременно на промените, които се случват в нея;
- про-активност - не просто действие в отговор на околната среда, а възможност за проява на целево поведение (проява на инициативност).

За нуждите на експерименталния модел в проекта се използват агенти от класа „обучаващи се агенти (Learning Agents)“. При тях обучението им дава възможност да работят самостоятелно в първоначално неизвестна среда и да станат по-компетентни от първоначалните си познания. Обучаващите се агенти се състоят от четири концептуални елемента (фиг. 2):

- обучаващ елемент (Learning Element) - отговаря за извършването на подобрения и надграждания, използва обратната връзка от критика за начина, по който агента се справя, и определя как трябва да бъде променен изпълняващия елемент, за да се справи по-добре в бъдеще;



Фиг. 2. Обучаващи се агенти

- изпълняващ елемент (Performance Element) - отговаря за избора на реактивни дейности;
- критик (Critic) - указва на обучаващият елемент колко добре агентът се справя по отношение на фиксиран стандарт за изпълнение; този елемент, оценяващ успеваемостта, е външен за агента, той не бива да го променя, за да нагажда индикациите му към собственото си поведение;
- генератор на задачи (Problem Generator) – отговорен за предлагане на действия, които ще доведат до натрупването на нов и информативен опит.

Предложеният модел е мулти-агентно-базирана софтуерна рамка, работата на която е разделена на четири слоя: мрежа, системен хардуер, транспорт на данни и системен софтуер. Тази рамка се състои от две части: Network Prevention (NP) и Host Prevention (HP). NP-компонентът функционира върху мрежовия интерфейс, който се явява външен за сървъра. Целта на HP-компонента е да защити дейностите на системата в режима на нейното ядро и от действия на потребителя (от вътрешни заплахи). HP-компонентът осъществява функции по наблюдение на критичните системни и потребителски дейности, които могат да бъдат потенциална мишена за хакерите в ядрото и за потребителя.

Моделът и експериментите показват следните функции и възможности на системата при защитата срещу заплахи в мрежата:

- защита срещу атаки и зловреден код – взаимодействието между NP и HP вътре в предложената система осигурява проактивна защита на крайните точки от атаки и зловреден код, в съответствие с изискванията на електронния бизнес;
- значително намаляване на неверните положителни и фалшивите отрицателни събития –

прилагането на поведенчески техники за анализ на данните в основата на системата и профилите за нормално поведение елиминират голяма част от неверните положителни и фалшивите отрицателни аларми;

- действия в реално време, висока производителност и гъвкавост;
- откриване на заплахи в ранен стадий – прилагането на стратегия за защита в дълбочина позволява откриването и блокирането на заплахи в ранен стадий;
- лесно изграждане на профил – използването на поведенчески техники за анализ на данните елиминира нуждата от използването на голямо количество от информация за обучение и изграждане на нормални профили, както е при техниките, базирани на аномалии;
- няма нужда от база данни със сигнатури на различните видове вируси, троянски коне и зловреден код – NP-компонентът има възможност да записва и обновява информацията автоматично, в зависимост от резултата от анализа. Допълнително системните администратори могат да допълват правилата за наблюдение и анализ;
- регистриране и записване на събитията в реално време на сървъра, за да може на по-късен етап експерти да анализират характеристиките на аномалното поведение на мрежовия трафик.

Тестването е проведено на база на разработени сценарии, с цел верификация и оценка на ефективността на архитектурата. Успешно са проведени експерименти като база за оценка на отделните компоненти и на цялата платформа. Тези експерименти удостоверяват, че системата отговаря на изисквания на съответните спецификации.

Резултатите показват, че предложената система работи достатъчно успешно при откриване на атаките и зловредния код, които са насочени към защитаваната система, с висока степен на точност и в реално време. Компонентът NP успява да характеризира нормалното поведение на TCP\IP протокола и да се открие най-простите атаки, целящи да засегнат заглавната част на пакетите. Компонент HP доказва високата си способност при защита срещу зловреден код, който влияе на операционни системи Windows, независимо дали зловредният код е в ядрото или насочен към потребителската дейност.

Заключение

Представеният в доклада експериментален модел за пореден път показва перспективите на използването на методите на „изкуствения интелект“ в системите за мрежова и информационна сигурност. Потвърждават се предварителните анализи на тези методи, които показват, че повечето от тях притежават както достойнства, така и недостатъци.

Една от основните цели на изпълнявания проект в следващите му раздели е, след като се проведе задълбочен анализ на експерименталното изследване на най-перспективните методи, да се направи опит за подобряване на положителните страни и избягване на недостатъците – или чрез модифициране на избрания метод, или чрез хибридизация на няколко избрани метода.

Благодарности

Научните изследвания, резултат от които са представени в настоящата публикация, са финансирани за сметка на договор Д 07/4 от 15.12.2016 г. към Фонд Научни изследвания по проект „Интелигентни методи за повишаване нивото на мрежовата и информационна сигурност”.

ЛИТЕРАТУРА

[1] The Next Generation of Cybercrime: How it’s evolved, where it’s going , SecureWorks. 2010

[2] The New Phishing Threat: Phishing Attacks A Proofpoint White Paper , Proofpoint, 2011

[3] Security in Telecommunications and Information Technology , ITU-T, June 2006

[4] Overview of cybersecurity Recommendation ITU-T X.1205 , April 2008

[5] Security intelligence for a faster world , Hewlett Packard, 2014

[6] IT Executive Guide to Security Intelligence IBM, January 2013

[7] Guide to Intrusion Detection and Prevention Systems (IDPS) NIST, Special Publication 800-94, February 2007

[8] Michael Luck, Peter McBurney, Christ Preist Agent Technology: Next Generation Computing AgentLink II, January 2003

[9] H. Albag, "Network & Agent Based Intrusion Detection Systems," https://www7.informatik.tu-muenchen.de/um/courses/seminar/worm/WS0405/albag.pdf

[10] S. D. Chi, J.S. Park, K.C. Jung and J.S. Lee Network Security Modeling and Cyber Attack Simulation Methodology Lecture Notes in Computer Science, Vol. 2119, 2001

[11] D. Dashgupta and F. Gonzales An Intelligent Intrusion Detection System Lecture Notes in Computer Science, Vol. 2052, 2001