Олексюк Б. Ю., магістрант, гр. КБм-21-1
Єфіменко А. А., к.т.н., доц.,
завідувач кафедри комп'ютерної інженерії та кібербезпеки
Вакалюк Т. А., д.пед.н., проф.,
професор кафедри інженерії програмного забезпечення
Державний університет «Житомирська політехніка»

МЕТОДИ ВИЯВЛЕННЯ ВТОРГНЕНЬ З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ

Методи виявлення вторгнень поділяються на дві групи, і існує кілька алгоритмів, які описані для навчання з наглядом і без нагляду.[1]

Алгоритми навчання з наглядом:

1. **k-найближчий сусід** – k-найближчий сусід – це класичний алгоритм, який знаходить k прикладів у навчальних даних, які є найближчими до тестового прикладу, і призначає їх найчастішу мітку серед цих прикладів новому прикладу. Єдиним вільним параметром є розмір k району.
2. **Багатошаровий персептрон** – навчання багатошарового персептрона включає оптимізацію ваг для функції активації нейронів, організованих у мережевій архітектурі. Глобальна цільова функція мінімізується за допомогою алгоритму RPROP. Вільний параметр — кількість прихованих нейронів. [2]
3. **Регулярний дискримінантний аналіз**. Припускаючи, що обидва класи прикладів розподілені нормально, байєсівською оптимальною розділювальною поверхнею є гіперплощина (LDA), якщо коваріаційні матриці однакові, або квадратична поверхня в іншому випадку (QDA). Поступове перетворення між двома випадками може бути реалізовано за допомогою параметра регуляризації . Інший вільний параметр ¸ контролює додавання ідентичної матриці до коваріаційних матриць.[3]
4. **Лінійна дискримінація Фішера** – Лінійна дискримінація Фішера будує розділову гіперплощину, використовуючи напрям, який максимізує міжкласову дисперсію та мінімізує дисперсію всередині класу для проекції навчальних точок на цей напрям. Вільний параметр – це компроміс між нормою напряму та «строгою» проекції.
5. **Машина лінійного програмування та допоміжна VectorMachine** – Машина лінійного програмування (LPM) і машина опорного вектора (SVM) конструюють гіперплощину мінімальної норми, яка розділяє два класи навчальних прикладів . LPM використовує 1-норму, SVMus — 2-норму. Крім того, SVM застосовує нелінійне відображення для побудови гіперплощини в просторі ознак. У наших експериментах використовуються радіальні базисні функції, їх складність контролюється параметром ширини w. Інший параметр C контролює компроміс між цією нормою гіперплощини і точність поділу.[3]

Алгоритми навчання без нагляду:

1. **Кластеризація k-середніх** – кластеризація k-середніх є класичним алгоритмом кластеризації. Після початкового випадкового присвоєння прикладу k кластерам обчислюються центри кластерів і приклади приписуються кластерам з найближчими центрами. Процес повторюється до тих пір, поки центри кластерів істотно не зміняться. Після того, як кластерне призначення фіксується, середня відстань прикладу до центрів кластера використовується як оцінка. Вільний параметр k.[1]
2. **Кластеризація за одинарним зв'язком** – кластеризація по одному зв'язку схожа на кластеризацію k-середніх, за винятком того, що кількість кластерів контролюється параметром distance W: якщо відстань від прикладу до найближчого центру кластера перевищує W, встановлюється новий кластер.[4]
3. **Машина опорних векторів чверті сфери** – SVM четвертої сфери є методом виявлення аномалій, заснованим на ідеї підгонки сфери до центру маси даних. Оцінка аномалії визначається відстанню точки даних від центру сфери. Вибір порогового значення для оцінок атаки визначає радіус сфери, що охоплює нормальні точки даних.[5]

Список використаних джерел

1. Pormohseni, Review and identify the computer Networkintrusion detection systems, 2011 (Language in Persian)
2. S Portnoy, L., Eskin, E., Stolfo, S.: Intrusion detection withunlabeled data using clustering. In: Proc. ACM CSSWorkshop on Data Mining Applied to Security. (2001)
3. SchÄolkopf, B., Smola, A.: Learning with Kernels. MITPress, Cambridge, MA (2002)
4. Mr. Manish Jain, Prof. Vineet Richariya "An Improved Techniques Based on Naïve Bayesian for Attack Detection" International Journal of Emerging Technology and Advanced Engineering Website www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 1, January 2012)

Paul Dokas, Levent Ertoz, Vipin Kumar, AleksandarLazarevic, Jaideep Srivastava, Pang-Nig Tan" DataMining for Network Intrusion Detection" Computer Science Department, 200 Union Street SE, 4-192 EE/CSC Building University of Minnesota, Minneapolis,MN 55455, USA .