# شناسایی تهدیدات درشبکه‌های رایانه‌ای با الگوریتم شاهین هریس

**هادی نیک‌منش*[۱]**

۱- کارشناس ارشد شبکه‌های کامپیوتری، مؤسسه غیرانتفاعی اقبال لاهوری مشهدhadiynikmanesh@gmail.com

**چکیده:**

در مقابله با حملات و نفوذ در شبکه‌های رایانه‌ای و اینترنت اشیاء از ابزارهای متعددی استفاده می‌شود که یکی از آنها سیستم تشخیص نفوذ است. این سیستم از چندین روش برای تشخیص نفوذ استفاده می‌کند که یکی از مهم‌ترین آنها روش تشخیص ناهنجاری است که امروزه بسیار مورد توجه قرار گرفته است. مشکل اساسی روش تشخیص ناهنجاری، نرخ هشدار غلط بالای آن است. تاکنون روش‌های مختلفی به‌منظور کاهش نرخ هشدار غلط و افزایشدقّت و نرخ تشخیص صحیح، در تشخیص نفوذ مورد استفاده و آزمایش قرارگرفته است؛ امّا هیچ یک نتوانسته‌اند به موفقیت کامل دست‌یابند. در این پژوهش، یک سیستم تشخیص‌ناهنجاری در شبکه‌های رایانه‌ای و اینترنت اشیاء با استفاده از الگوریتم بهینه‌سازی شاهین هریس ارایه شده است. مجموعه ویژگی انتخاب‌شده توسط الگوریتم پیشنهادی، بیش‌ترین تأثیر را در تشخیص ناهنجاری تراکنش‌ها دارد. استفاده از الگوریتم شاهین هریس به همراه طبقه‌بندی درخت تصمیم، میزان دقّت را به ۹۹,۸۳۳ درصد افزایش داده است و در مقایسه با الگوریتم‌های مشابه نشان از بهبودآن دارد.

**کلید واژه‌ها:** اینترنت اشیاء، تشخیص ناهنجاری، الگوریتم درخت تصمیم.

## Identification of computer network threats with Shahin Harris algorithm

**Hadi Nik Manesh*[1]**

1-Senior expert in computer networks, Iqbal Lahori non-profit Institute, Mashhad, hadiynikmanesh@gmail.com

**Abstract:**

In dealing with attacks and intrusions in computer networks and Internet of Things, several tools are used, one of which is the intrusion detection system. This system uses several methods for intrusion detection, one of the most important of which is the anomaly detection method, which is of great interest today. The main problem of the anomaly detection method is its high false alarm rate. So far, various methods have been used and tested in intrusion detection in order to reduce the false alarm rate and increase the accuracy and correct detection rate; but none of them have been able to achieve complete success. In this research, an anomaly detection system in computer networks and Internet of Things is presented using Shahin Harris optimization algorithm. The set of features selected by the proposed algorithm has the greatest effect in identifying the abnormality of transactions. The use of Shahin Harris algorithm along with the decision tree classifier increased the accuracy to 99.833% and shows an improvement compared to similar algorithms.

**Keyword**: Internet of things, anomaly detection, decision tree algorithm.

## ۱- مقدمه

مفهوم اینترنت اشیا[1] IoT در سال ۱۹۹۹م [۱] بعد از ورود دستگاه‌های بی‌سیم به بازار و معرفی فناوری‌های RFID[2] و شبکه‌های سنسور بی سیم[3][۲] معرفی شد. اینترنت اشیاء یکی از رایج‌ترین فناوری‌هایی است که امکان اتصال داخلی بین دستگاه‌ها یا اشیای انبوه را از طریق اینترنتیا اینترانت فراهم می‌سازد و این امکان را به اشیاء می‌دهد تا با یکدیگر ارتباط برقرار سازند[۳].

سیستم‌های تشخیص نفوذ، یک زمینه‌ی تحقیقاتی مهم در زمینه‌ی امنیّت شبکه‌های اینترنت اشیااست. سیستم تشخیص نفوذ، سخت‌افزار و یا نرم‌افزاری است که کار نظارت بر شبکه‌های کامپیوتری را در مورد فعالیت‌های مخرب و یا نقض سیاست‌های مدیریتی و امنیتی انجام می‌دهد و گزارش‌های حاصله را به بخش مدیریت شبکه ارائه می‌دهد[۴].

از زمان ارائه‌ی سیستم‌های تشخیص نفوذ تا به امروز، این سیستم‌ها با مشکلات بسیاری روبرو بوده‌اند که با تحقیقات انجام‌شده و بهبودهای صورت‌گرفته در جهت رفع مشکلات آن‌ها، روزبه‌روز به سمت تکامل خود پیش می‌روند. یکی از مشکلات اساسی که سیستم‌های تشخیص نفوذ با آن مواجه هستند، تعداد زیاد هشدارها و بالا بودن نرخ تشخیص نادرست در بین هشدارها است.

در سال‌های اخیر از الگوریتم‌ها و روش‌های هوش مصنوعی مختلفی در طراحی سیستم‌های تشخیص نفوذ استفاده شده است. چالش‌هایی که هم‌چنان در استفاده از این الگوریتم‌ها وجود دارد، انتخاب بهترین مجموعه از ویژگی‌ها برای داده‌های مورد استفاده و انتخاب بهترین پارامترها برای روش‌های طبقه‌بندی است. برای رفع این مشکل، الگوریتم‌های فراابتکاری پنجره جدیدی به روی محققان گشوده است. در مطالعات پیشین از روش‌های تکاملی با اهداف مختلفی بهره گرفته شده است. به طور مثال، رایس و همکاران، از روش الگوریتم مورچگان (ACO) به منظور انتخاب بهترین ویژگی‌ها در حل مسأله‌ی تشخیص نفوذ به شبکه بهره گرفته شده است[۵]. علی و همکاران از الگوریتم ازدحام ذرات (PSO) به منظور انتخاب وزن‌های بهینه در شبکه‌های یادگیری سریع (FLN) در کاربرد IDS استفاده کرده‌اند[۶]. یانگ و همکاران از الگوریتم کلونی زنبورها(ABC) با هدف بهینه کردن پارامترهای طبقه‌بندی شبکه‌های بیزین بهره گرفتند و از آن در تشخیص نفوذ به شبکه استفاده کرده‌اند[۷]. با مروری بر سایر پژوهش‌های انجام شده‌می‌توان نتیجه گرفت که الگوریتم‌های فراابتکاری به صورت گسترده‌ای در حل مسائل تشخیص نفوذ به شبکه و در یکی از دو کاربرد انتخاب پارامترهای بهینه‌ی روش‌های طبقه‌بندی و یا انتخاب ویژگی‌های بهینه به کار گرفته شده‌اند.

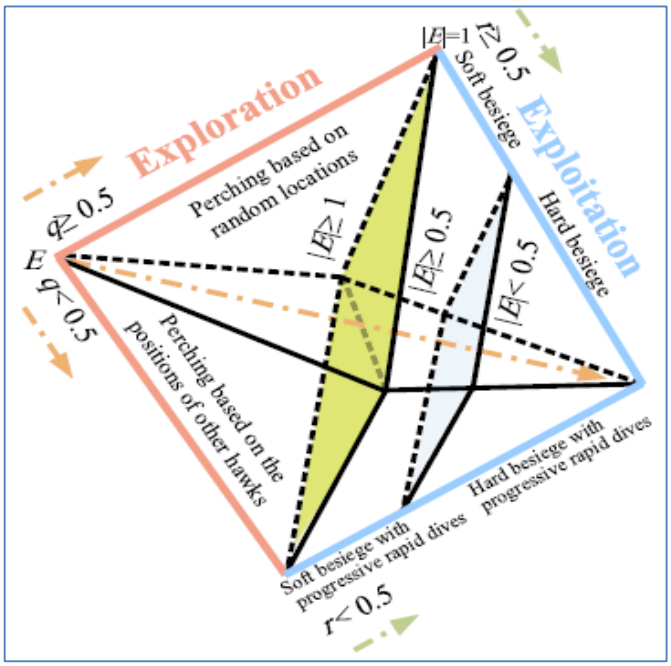
از آن جایی‌که سرعت و دقّت در طراحی سیستم‌های تشخیص نفوذ از اهمیت بالایی برخوردار است، الگوریتم بهینه‌سازی شاهین هریس، به عنوان یک الگوریتم فراابتکاری که از قدرت همگرایی بالا و سرعت زیادی برخوردار است، گزینه مناسبی به نظر می‌رسد. در این مقاله، برای تشخیص ناهنجاری در شبکه از روش‌های مُبتنی بر دسته‌بندی (درخت تصمیم و دسته‌بندی بیزین) و مُبتنی بر خوشه‌بندی (روش نزدیک‌ترین هم‌سایه) و برای انتخاب ویژگی‌ها از الگوریتم شاهین هریس استفاده می‌شود.

## ۲- الگوریتم شاهین هریس

در این پژوهش، الگوی بهینه‌سازی مُبتنی بر جمعیت و الهام از طبیعت، ارائه شده است که به آن بهینه‌ساز هریس هاوکس (HHO) یا شاهین هریس گفته می‌شود. الگوریتم HHO از رفتار مشارکتی و سبک تعقیب طعمه توسط گروه شاهین هریس در طبیعت با روش گشت‌وگذار غافل‌گیرانه استنباط شده است. در این استراتژی هوشمند، چندین

[1] Internet of Thing
[2] Radio-frequency identification
[3] Wireless Sensor Networks(WSN)

شـــاهین با حمله از جهات مختلف، تلاش دارند طعمه (خرگوش) را غافلگیر کنند. هر الگوریتم تکاملی جدیدی که معرفی می‌شـــود برای اثبات قدرت خود در بهینه‌ســـازی باید بر روی دو فاز مهم یعنی اکتشـــاف (Exploration) و بهره‌برداری (Exploitation) تمرکزداشـــته باشـــد و بتواند تعادل خوبی بین این دو فاز ایجاد کند. در فاز اکتشـــاف، الگوریتم به‌دنبال کشف فضای جستجو است؛ و به‌روزرسانی موقعیت شاهین‌ها، بر اساس عملگر تصادفی و یا بر اساس موقعیت سایر شاهین‌ها صورت می‌گیرد و در فاز بهره‌برداری، الگوریتم به دنبال بهبود جواب‌های پیدا شده ا ست و از عملگرهای محاصره سخت و محاصره نرم استفاده می‌کند. در شکل ۱ فازهای مختلف الگوریتم دیده می‌شود[۸].




شکل (۱): فازهای مختلف الگوریتم شاهین[۸].

**۱-۲- فاز اکتشاف**

در HHO، شـــاهین‌های هریس به طور تصـــادفی در موقعیت‌های مختلفی (مکان‌های بلند) قرار می‌گیرند و منتظر می‌مانند تا طعمه‌ای را بر ا ساس دو ا ستراتژی شنا سایی کنند. اگر شانس q را برای هر ا ستراتژی قرارگیری در نظر بگیریم و شـــاهین‌ها با توجه به موقعیت دیگر اعضـــای خانواده خود (تا هنگام حمله به اندازه‌ی کافی نزدیک به آنها با شند) و موقعیت طعمه یا خرگوش قرار بگیرند، موقعیت آنها با معادله ۱ و شرط $q < 0.5$ مدل می شوند و اگر روی درختان بلند به طور تصادفی (در محدوده‌ی خانه‌ی گروه) قرار بگیرند، با شرط $q \geq 0.5$ مدل می‌شوند[۸].

$$X(t+1) = \begin{cases} X_{rand}(t) - r_1|X_{rand}(t) - 2r_2X(t)| & q \geq 0.5 \\ \left(X_{rabbit}(t) - X_m(t)\right) - r_3\left(LB + r_4(UB - LB)\right) & q < 0.5 \end{cases} \quad (۱)$$

در این معادله، $X(t+1)$ بردار موقعیت شـــاهین ها در تکرار بعدی t، $X_{rabbit}$ (t) موقعیت خرگوش، $X(t)$ بردار موقعیت فعلی شاهین‌ها، $r_1, r_2, r_3, r_4$ و q اعداد ت صادفی بین صفر و یک ه ستند که در هر تکرار به‌روز می شوند، LB و UB مرزهای بالا و پایین متغیرها (محدوده خانه گروه) را نشـــان می‌دهد، $X_{rand}(t)$ موقعیت تصـــادفی یک شـــاهین

انتخاب شده از جمعیت فعلی ا ست و $X_m$ موقعیت متو سط جمعیت فعلی شاهین‌ها ا ست که از رابطه‌ی زیر به د ست می‌آید:

$$X_m(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t) \quad (۲)$$

که در آن $X_i(t)$ موقعیت هر شاهین را در تکرار t و N تعداد کل شاهین‌ها را نشان می‌دهد[۸].

**۲-۲- فاز بهره‌برداری**

با توجه به رفتارهای گریز طعمه و راهبردهای تعقیب و گریز شـــاهین هریس، چهار مدل اســـتراتژی احتمالی در HHO برای مدل‌سازی مرحله حمله پیشنهاد شده است.

فرض کنید که r شـــانس طعمه در فرار موفق ($r < 0.5$) یا فرار ناموفق ($r \geq 0.5$) قبل از گشـــت‌وگذار غافل‌گیرانه است. شاهین‌ها بسته به انرژی احتمالی طعمه، طعمه را از جهات مختلف محاصره می‌کنند. در شرایط واقعی، شاهین‌ها به طعمه موردنظر نزدیک‌تر و نزدیک‌تر می‌شـــوند تا با انجام گشـــت‌وگذار غافل‌گیرانه، شـــانس خود را در گرفتن طعمه افزایش دهند. پس از چند دقیقه که طعمه انرژی بیشـــتری را از دســـت می‌دهد، شـــاهین‌ها روند محاصـــره را تشـــدید می‌کنند تا طعمه گیج‌شده را گرفتار کنند. برای مدل‌سازی این استراتژی، از پارامتر E استفاده می‌شود. در این رابطه، هنگامی که $|E| \geq 0.5$، محاصره نرم اتفاق می‌افتد، و زمانی که $|E| < 0.5$، محاصره سخت رخمی دهد[۸].

**۲-۲-۱- محاصره نرم**

هنگامی که $r \geq 0.5$ و $|E| \geq 0.5$، طعمه هنوز هم به اندازه‌ی کافی انرژی دارد و ســـعی در فرار با برخی جهش‌های گمراه‌کننده تصـــادفی دارد، اما در نهایت نمی‌تواند فرار کند. در طی این تلاش‌ها، شـــاهین‌های هریس آن را محاصـــره می‌کنند تا خرگوش را خسته‌تر کنند. این رفتار طبق معادلات زیر انجام می‌شود:

$$X(t+1) = \Delta X(t) - E|JX_{rabbit}(t) - X(t)| \quad (۳)$$

$$\Delta X(t) = X_{rabbit}(t) - X(t) \quad (۴)$$

در جایی که $\Delta X(t)$ تفاوت بین بردار موقعیت خرگوش و موقعیت فعلی شـــاهین در تکرار t اســـت، $r_5$ یک عدد تصادفی بین صفر و یک و $J = 2(1 - r_5)$ نشان‌دهنده قدرت پرش تصادفی خرگوش هنگام فرار است. مقدار J به طور تصادفی در هر تکرار تغییر می‌کند تا ماهیت حرکات خرگوش را شبیه‌سازی کند[۸].

**۲-۲-۲- محاصره سخت**

هنگامی‌که $r \geq 0.5$ و $|E| < 0.5$، طعمه بسیار خسته است و انرژی فرار کمی دارد. علاوه بر این، شاهین‌های هریس به ســـختی طعمه موردنظر را محاصـــره می‌کنند تا در نهایت گشـــت‌وگذار غافل‌گیرانه را انجام دهند. در این شـــرایط موقعیت‌های فعلی با استفاده از معادله زیر به‌روز می‌شوند:

$$X(t+1) = X_{rabbit}(t) - E|\Delta X(t)| \quad (۵)$$

**۲-۲-۳-محاصره نرم با شیرجه‌های سریع مترقی**

هنگامی که $r < 0.5$ و $|E| \geq 0.5$، خرگوش دارای انرژی کافی برای فرار موفق اســـت و هنوز هم محاصـــره نرم قبل از گشـــت‌وگذار غافل‌گیرانه انجام می‌شـــود. این رویه نســـبت به مورد قبلی هوشـــمندانه‌تر اســـت. برای مدل‌ســـازی ریاضـــی

الگوهای فرار طعمه و حرکات جهشی، از مفهوم LF در الگوریتم HHO استفاده می‌شود. LF برای تقلید از حرکات فریبنده زیگزاگی طعمه در مرحله فرار و شیرجه‌های نامنظم، ناگهانی و سریع شاهین در اطراف طعمه استفاده می‌شود. در واقع، شاهین‌ها چندین شیرجه سریع تیمی را در اطراف خرگوش انجام می‌دهند و سعی می‌کنند با توجه به حرکات فریبنده‌ی طعمه، مکان و مسیرهای خود را به تدریج تصحیح کنند. با الهام از رفتارهای واقعی شاهین‌ها، فرض شده است که برای قرار دادن طعمه در موقعیت‌های رقابتی، شاهین‌ها می‌توانند بهترین شیرجه را به سمت طعمه انتخاب کنند. بنابراین، برای انجام یک محاصره نرم، فرض می‌شود که شاهین‌ها می‌توانند حرکت بعدی خود را بر اساس معادله زیر ارزیابی کنند:

$$Y = X_{rabbit}(t) - E|JX_{rabbit}(t) - X(t)| \quad (۶)$$

سپس، آن‌ها نتیجه‌ی احتمالی چنین حرکتی را با شیرجه‌ی قبلی مقایسه می‌کنند تا دریابند که این یک شیرجه‌ی خوب خواهد بود یا نه. اگر بینند طعمه حرکات فریبنده‌تر را انجام می‌دهد، شروع به انجام شیرجه‌های نامنظم، ناگهانی و سریع می‌کنند. ما تصور می‌کنیم که آنها با استفاده از معادله زیر بر اساس الگوهای LF شیرجه بزنند:

$$Z = Y + S \times LF(D) \quad (۷)$$

جایی که D ابعاد مسأله است و S یک بُردار تصادفی با ابعاد $1 \times D$ و LF تابعی است که با استفاده از معادله زیر محاسبه می‌شود:

$$LF(x) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left( \frac{L(1+\beta) \times \sin\frac{\pi\beta}{2}}{L(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right) \quad (۸)$$

در جایی که u، v مقادیر تصادفی بین صفر و یک هستند، β یک مقدار ثابت پیش‌فرض برابر ۱,۵ است. از این رو، استراتژی نهایی برای به‌روزرسانی موقعیت شاهین‌ها در فاز محاصره نرم توسط معادله زیر قابل اجرا است:

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \quad (۹)$$

جایی که Y و Z با استفاده از معادلات ۶ و ۷ به دست می‌آیند.

**۲-۲-۴- محاصره سخت با شیرجه‌های سریع مترقی**

وقتی $r < 0.5$ و $|E| < 0.5$ ، خرگوش انرژی کافی برای فرار ندارد و محاصره سخت قبل از گشت‌وگذار غافل‌گیرانه برای گرفتن و کشتن طعمه شکل می‌گیرد. وضعیت این مرحله در سمت طعمه شبیه به شرایط موجود در محاصره نرم است، امّا این بار شاهین‌ها سعی می‌کنند متوسط فاصله خود را با طعمه کاهش دهند. بنابراین، قانون زیر در شرایط محاصره سخت انجام می‌شود:

$$X(t+1) = \begin{cases} Y & \text{if } F(Y) < F(X(t)) \\ Z & \text{if } F(Z) < F(X(t)) \end{cases} \quad (۱۰)$$

جایی که Y و Z با استفاده از قوانین جدید در معادلات زیر به دست می‌آیند.

$$Y = X_{rabbit}(t) - E|JX_{rabbit}(t) - X_m(t)| \quad (۱۱)$$

$$Z = Y + S \times LF(D) \quad (۱۲)$$

جایی که $X_m(t)$ با استفاده از معادله ۲ به دست می‌آید.

شبه‌الگوریتم پیشنهادی HHO در الگوریتم ۱ (شکل ۲) گزارش شده است[۸].

---

levy flight

**الگوریتم ۱. شبه کد الگوریتم شاهین هریس(HHO)**

**ورودی‌ها:** اندازه جمعیت N و حداکثر تعداد تکرارT

**خروجی‌ها:** موقعیت خرگوش و مقدار شایستگی آن

ساخت جمعیت تصادفی $X_i(i = 1, 2, \ldots, N)$

**تا زمانی که** (شرط توقف برآورده نشده است) مراحل زیر را **انجام دهید:**

-محاسبه مقادیر شایستگی شاهین ها

-تعیین $X_{rabbit}$ به عنوان موقعیت خرگوش (بهترین جواب)

**برای** (هر شاهین $X_i$ ) مراحل زیر را **انجام دهید:**

بروز کردن انرژی اولیه $E_0$ و قدرت پرش J

$J = 2(1-rand\ ()) \ , E_0 = 2rand\ () - 1$

بروز کردن E با استفاده از معادله ۳-۵

**اگر** ($|E|\geq 1$) **سپس** (مرحله اکتشاف: بروز کردن بردار مکانی با استفاده از معادله ۳-۳)

**اگر** ($|E|<1$) **سپس** (مرحله بهره برداری)

**اگر** ($r\geq 0,5$ و $|E|\geq 0,5$) سپس (محاصره نرم: بروز کردن بردار مکانی با استفاده از معادله ۳-۶)

**در غیر این صورت اگر** ($r\geq 0,5$ و $|E|<0,5$) سپس

(محاصره سخت: بروز کردن بردار مکانی با استفاده از معادله ۳-۸)

**در غیر این صورت اگر** ($r<0,5$ و $|E|\geq 0,5$) سپس

(محاصره نرم با شیرجه‌های سریع مترقی: بروز کردن بردار مکانی با استفاده از معادله ۳-۱۲)

**در غیر این صورت اگر** ($r<0,5$ و $|E|<0,5$) سپس

(محاصره سخت با شیرجه‌های سریع مترقی: بروز کردن بردار مکانی با استفاده از معادله ۳-۱۳)

**برگرداندن** $X_{rabbit}$

شکل (۲): شبه‌کُد الگوریتم شاهین هریس.

## ۵- روش پیشنهادی

در الگوریتم پیشنهادشده، تا نمونه‌ای تصادفی و بدون جای‌گذاری به اندازه‌ی ۱۰۰۰۰ را از مجموعه‌ی داده Kddcup99 انتخاب شود و با استفاده از الگوریتم شاهین هریس با تعداد جمعیت ۳۰ و با ماکزیمم تکرار ۲۰ به انتخاب ویژگی پرداخته و این الگوریتم را با الگوریتم ژنتیک و گرگ خاکستری مقایسه کرده‌ایم. سپس، از الگوریتم درخت تصمیم جهت طبقه‌بندی داده‌های آموزشی انتخاب ویژگی شده استفاده شده است. جهت ارزیابی، روش پیشنهادی خود را با روش‌های Navie Bayes و K-Nearest Neighbor مقایسه کرده و از لحاظ دقّت به ترتیب الگوریتم پیشنهادی در جایگاه نخست و سپس، الگوریتم نزدیک‌ترین همسایه و در مرتبه بعد، الگوریتم بیزین بهترین نتایج را به‌دست آورده‌اند. نتایج در قالب نمودارهایی در ادامه نشان داده شده است. در الگوریتمK-Nearest Neighbor مقدار k را برابر ۵ در نظر گرفته شده است. همان طور که در جدول ۱ مشاهده می‌کنید استفاده از الگوریتم شاهین نسبت به گرگ خاکستری و ژنتیک از دقّت و نرخ تشخیص بالاتری برخوردار است.

جدول (۱): نمایش پارامترهای مختلف ارزیابی با استفاده از روش‌های مختلف انتخاب ویژگی.

| پارامتر ارزیابی | HHO+DT | GWO+DT | GA+DT |
|---|---|---|---|
| Accuracy (دقت) | ۹۹,۸۳۳ | ۹۹,۶۶۷ | ۹۹,۵۸ |
| DR (نرخ تشخیص) | ۱۰۰ | ۱۰۰ | ۹۶,۱۱ |

| | | | |
|---|---|---|---|
| FPR (نرخ مثبت کاذب) | ۰,۰۰۰۶۷۰۰۲ | ۰,۰۰۰۳۳۵۰۱ | ۰,۷ |

بعد از انتخاب ۱۰ ویژگی توسط الگوریتم شاهین هریس، از دسته‌بندهای مختلف نزدیک‌ترین همسایه، درخت تصمیم CART و دسته‌بندی بیزین، برای یافتن بهترین دسته‌بندی استفاده شده است. همان‌گونه که در جدول ۲ مشاهده می‌شود روش پیشنهادی (استفاده از الگوریتم CART) از نرخ هشدار کاذب قابل قبولی برخوردار است.

جدول (۲): مقایسه نتایج روش‌ها.

| پارامتر ارزیابی | HHO+CART* | HHO+KNN | HHO+NB |
|---|---|---|---|
| Accuracy (دقت) | ۹۹,۸۳۳ | ۹۹,۷ | ۹۸,۱۳۳ |
| Precision (صحّت) | ۸۸,۲۳۵ | ۱۰۰ | ۱۰۰ |
| DR (نرخ تشخیص) | ۱۰۰ | ۱۰۰ | ۸۸,۸۸۹ |
| FPR (نرخ مثبت کاذب) | ۰,۰۰۰۶۷۰۰۲ | ۰ | ۰ |
| Specificity (خاصیت) | ۹۹,۹۳۳ | ۱۰۰ | ۱۰۰ |

## ۶- نتیجه‌گیری و جمع‌بندی

در این پژوهش، شناسایی و انتخاب ویژگی‌های تأثیرگذار در تعیین وضعیت نرمال یا ناهنجاری تراکنش‌ها مورد ارزیابی قرار گرفته است. کاهش ویژگی‌ها، برای بهبود عملکرد در زمان یادگیری، دقّت دسته‌بندی و به دست آوردن قوانین قابل‌فهم، ضروری می‌باشد؛ البته کاهش ویژگی‌ها نباید موجب کاهش نرخ تشخیص و افزایش نرخ تشخیص کاذب شود. بنابراین، نیاز به الگوریتمی داریم که زیرمجموعه قدرتمندی از ویژگی‌ها را برایما به ارمغان داشته باشد، در این زمینه از الگوریتم شاهین استفاده کردیم و نشان دادیم که استفاده از این الگوریتم نسبت به گرگ خاکستری و ژنتیک از دقّت و نرخ تشخیص بالاتری برخوردار است. نتایج حاکی از آن بود که بهترین ویژگی‌ها در مجموعه داده نمونه تشخیص نفوذ، تعداد ۱۰ ویژگی بود. این مجموعه ویژگی، بیش‌ترین تأثیر را در تعیین نرمال بودن یا ناهنجار بودن تراکنش‌ها دارا بودند.

روش پیشنهادی در این پایان‌نامه دارای نرخ تشخیص بسیار مطلوب و نرخ هشدار غلط پایین است. نرخ هشدار غلط بالا سبب می‌شود که بار محاسباتی زیادی به سیستم تحمیل شده و سبب کندی و ناکارآمدی آن در تشخیص نفوذها و حملات گردد. موارد فوق حاکی از آن است که می‌توان از مدل طبقه‌بندی پیشنهادی برای سایر مجموعه داده‌های تشخیص نفوذ به منظور یافتن الگوی مناسب طبقه‌بندی تراکنش‌ها به دو طبقه نرمال و ناهنجاری استفاده نمود.

## منابع:

[1] Gubbi, J. Buyya, R. Marusic, S. and Palaniswami, M., ‘‘Internet of things (IoTs): a vision, architectural elements, and future directions,’’ Future Generation Computer Systems, Vol. 29, No. 7, pp. 1645–1660, 2013.

[2] Riahi Sfar, A. Natalizio, E. Challal, Y. and Chtourou, Z., ‘‘a roadmap for security challenges in the Internet of Things,’’ Digital Communications and Networks 4, pp. 118–137, 2018.
[3] Deepa V. Jose, Vijyalakshmi A., ‘‘an Overview of Security in Internet of Things,’’ Procedia Computer Science, 143, pp. 744–748, 2018.
[4] Alampalayam Kumar, S. Vealey, T. and Srivastava, H., ‘‘Security in Internet of Things: Challenges, Solutions and Future Directions,’’ IEEE Computer Society 16, pp. 1530-1605, 2016.
[5] Rais, H. M. and Mehmood, T., ‘‘Dynamic Ant Colony System with Three Level Update Feature Selection for Intrusion Detection,’’ IJ Network Security, 20, pp. 184-192, 2018.
[6] Ali, M., Al Mohammed, B. A. D., Ismail, A. and Zolkipli, M. F., ‘‘A new intrusion detection system based on fast learning network and particle swarm optimization,’’ IEEE Access, 6, pp. 20255-20261, 2018.
[7] Yang, J. Ye, Z. Yan, L. Gu, W. and Wang, R., ‘‘Modified Naive Bayes Algorithm for Network Intrusion Detection based on Artificial Bee Colony Algorithm,’’ International Conferences on Intelligent Data Acquisition and Advanced Computing Systems, pp. 35-40, 2018.
[8] Heidari, A. A. et al., ‘‘Harris hawks optimization: Algorithm and applications,’’ Future Generation Computer Systems, pp. 849–872, 2019.