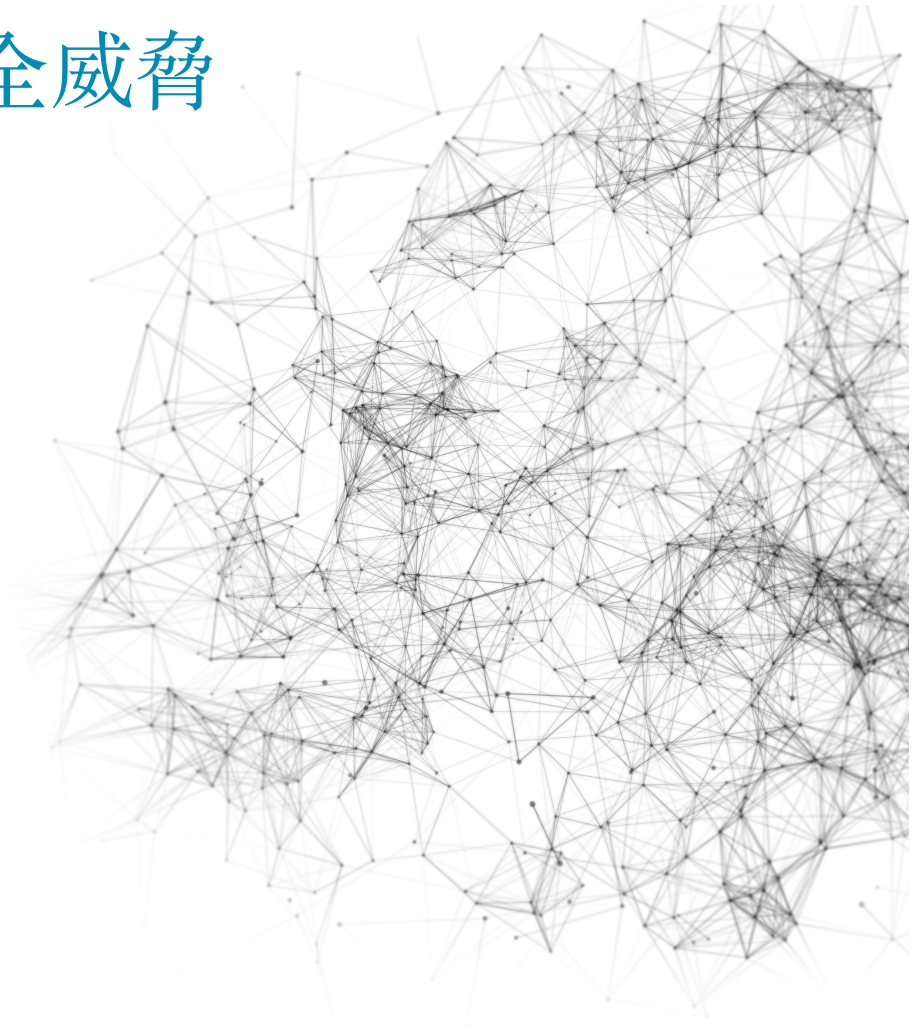
新一代DDoS安全防禦新思維

以機器學習 人工智慧有效防禦DDoS安全威脅



技術白皮書

GenieNetworks

突破性技術(Disruptive Technologies)新趨勢：人工智慧 機器學習

網路入侵與攻擊的方式眾多，其中分散式阻斷攻擊 (Distributed Denial of Service，以下簡稱DDoS攻擊) 是一種透過持續性的強大攻擊流量癱瘓目標網站的攻擊，使其對外服務中斷以影響信譽或營利，時常被利用來作為達成如財務勒索、政治性威脅抗爭、惡性商業競爭的工具。DDoS攻擊因為發動門檻低，配合著各種新興網路技術如5G、IoT等推波助瀾，還會因應網路特性採取最合適的攻擊網路技術，使得近年來不僅攻擊規模越來越大、攻擊次數也越來越頻繁，已然成為最令人頭痛的資安攻擊型態之一。

DDoS攻擊問世多年，但至今仍然是令聞之色變的手法，主要原因在於隨著新的網路及資訊技術的出現，其攻擊技巧也不斷地翻新、進化，對網路環境造成越來越劇烈的危害，其攻擊行為也更善於偽裝，以躲避資安技術的防堵。像是時下的科技新顯學：人工智慧(Artificial Intelligence，AI) 與機器學習(Machine Learning，ML)，也成了被駭客利用的攻擊工具進化器。

自從2016年AlphaGo打敗人類棋士，證明了AI 技術不再只是高等學術殿堂中的理論型研究，更能有效地應用於實務上，一夕之間人工智慧技術攻占了各大媒體版面。新一代的駭客，也開始開發基於AI和機器學習的尖端攻擊工具，生成特殊客製化的攻擊程式。這些攻擊不僅僅是基於靜態演算法的變體，更採用自動化和機器學習的技術將它們客製化到一個獨特的目標。藉由這樣的手法，不僅能擴大攻擊的範圍和規模，提高其攻擊行為的成功率，也同時使它們更難以被檢測。

其中的一個例子，是使用具自我學習技術的智慧攻擊設備集群，取代傳統的僵屍網路(Botnets)，以發動更有效的攻擊，像是發展蜂巢網路(Hivenets)和機器人集群(Swarmbots)的組合。它的概念上是讓許多小型惡意程式攻入有弱點的設備，並以昆蟲在蜂群或群居時所做的相同方式相互聯繫，以聰明地觀察活動、識別弱點、並共同決定何時何地進行攻擊。同時，「蜂巢」可以給一群「昆蟲」指令，以適應其環境的方式去實施，這個蜂巢可以根據當地情報的傳播和分享，來決定它的實施和時間安排，不需要中繼站指揮與控制，就可自主判斷執行持續性滲透。隨著Hivenet識別並攻陷更多的設備，它將能夠以指數級增長，不僅能同時攻擊多個受害者，擴大其攻擊能力及規模，也能阻礙防禦系統抑制與回應威脅的能力。

另一個例子，是將自然語言處理 (Natural Language Processing，NLP)演算法，用於分析公開或竊取的社群網路及郵件訊息等，再依這些個人化資訊發起針對性的攻擊，以提高網路釣魚(phishing) 類社交工程(social engineering)的攻擊成功率。

另一個駭客運用機器學習技術的例子是「生成對抗網路」(Generative Adversarial Network，GAN)。「生成對抗網路」的概念，就像是一個藝術贋品的製造者，不斷地人偽造世界名畫。他所生產的贋品，也不斷地交由名畫鑑定專家來鑑定真偽。而贋品的製造者會持續得到專家鑑定結果的反饋，且依據每次得到的反饋，再進一步改良其偽造技巧。透過如此不斷回饋與改良的過程，最後終將使得贋品真假難辨 (或至少是沒有任何一個鑑定專家能判斷真偽)。「生成對抗網路」架構的技術，也開始被攻擊者巧妙利用，以繞過資安防禦系統的檢測。

換句話說，隨著突破性新技術(Disruptive Technologies)如人工智慧及機器學習的興起和普及，越來越多的網路犯罪分子已開始利用新科技提供的新機會進行攻擊，為網路安全帶來前所未見的重大威脅。

科技無善惡，善惡在人心：以子之AI，攻子之AI

隨著網路技術的不斷演進，網路駭客利用人工智慧機器學習等新技術不斷翻新其攻擊技巧，不僅對網路環境造成越來越劇烈的危害，也更能躲避資安防護的防堵技術。

傳統的資安防禦技術，在面對這些不斷進化的自動化智慧型攻擊技巧會有幾項問題：首先，傳統的資安防禦常利用所謂的特徵比對式偵測(Signature-based detection)。這類的資安技術，不僅是要等到攻擊出現、對網路或使用者已經造成危害，還要再等到所謂的資安專家團隊對監測到的攻擊活動進行分析、產出所謂的「攻擊特徵碼/病毒碼」之後，才能據以對接下來同樣的攻擊活動進行反制。它的缺點在於攻擊特徵碼的產出因為需要專家人力的介入，不僅所需的時間較長，對攻擊活動所需的防堵反應時間也就較長，進而使得被攻擊的目標受到較大的衝擊和傷害。再者，運用人工智慧的新攻擊模式會自動地持續改變，專家人力資源有限，因而也疲於對大量多樣化的攻擊行為進行即時回應調整。

此外，面對更自動化且具智能的攻擊，傳統的資安防禦技術也會有檢測屬性過於單一平面(single vector)的疑慮。舉例來說，以攻擊病毒碼的檢測方式為例，它的檢測屬性是流量內容值比對；以黑名單IP位址比對的檢測方式為例，它的檢測屬性是流量來源或目地的IP位址；以流量行為基線比對的檢測方式為例，它的檢測屬性是符合特定流量行為的速率(bps, pps)或比率(ratio)……等等。由這些例子可知，這些傳統的資安防禦技術多半利用單一的流量特徵屬性去進行攻擊的判定。然而，惡意程式內容、IP 位址、和流量行為等是網路攻擊的基本組成部分，在人工智慧的輔助下，他們可以持續地自動被改變和調整，使他們更難被發現和判定，這會使得大多數採用人為決定性(deterministic)檢測屬性的傳統安全解決方案顯得過時。

再者，除了專家人力介入的需求與檢測屬性定義的自動化外，因為DDoS攻擊具有分散式攻擊來源的這項特性，也使得攻擊的檢測和防禦尤其困難。DDoS攻擊來源是全面性的，攻擊流量會由網路各地區而來，難以用單一位置部署、單一防禦設備來加以分析及判定攻擊流量。尤其，當需要防護的網路範圍廣大、網路組成架構複雜時，如電信運營商網路的全網DDoS防禦，面對分散又多變的攻擊行為，即便是專家也難以完全了解他轄內廣大的網路流量的動態成分與特性，進而能正確地定義攻擊流量的行為模式和檢測屬性。

如果將「資訊安全」裡攻擊駭客與資安防禦的對抗以戰爭作比喻，這場資安戰爭是一場不斷進化、永無止息的戰爭。在駭客學習AI、取得演算法、雲端計算能力等成本都越來越低的現今環境中，傳統的資安防禦技術顯得左支右絀。面對這些不斷進化的攻擊技巧，資安防禦方也秉持著追求「魔高一尺，道高一丈」的精神，持續發展更新的防禦技術，走向流量行為的自動智能分析，不僅減少對專家人力需求的仰賴，也能高效處理更複雜且動態的行為特徵分析，讓防禦端的偵測和回應速度更快及精準。因此人工智慧、機器學習等技術，也成為資安防禦技術中不可獲缺的重要角色。

近年由於硬體儲存成本的大幅下降和運算能力增強，加上巨量數據取得的可能性，使得今日的機器學習技術已能從資料中自行學習出規律，做為實現人工智慧的一種方式。一般談到機器學習的應用時，技術上的關鍵部分包含了

一、資料清整(Data Cleaning)：機器要從巨量資料中挖掘出規律，「乾淨」的數據在分析時是第一個關鍵；

二、特徵萃取(Feature Extraction)：特徵萃取是要從資料中挖出可以用的特徵，比如流量的來源位

址、速率、時間等；再把特徵量化，將每筆流量轉變成一個多維度的向量；

三、特徵選擇(Feature Selection)：根據檢視機器學習模型學習的結果，去選擇什麼樣的特稱是比較重要的。以分析攻擊流量為例，流量的來源分布、速率、時段變化…等等，可能就是比較重要的特徵，而其他向量的影響可能並不會那麼顯著。藉由逐步測試、或使用演算法篩選特徵，找出最恰當的特徵組合以讓學習的效果最佳化；

四、模型選取：根據所要解決的問題、擁有的資料類型和過適化(overfitting)等情況進行衡量評估，選擇性能合適的機器學習模型。一般而言，機器學習模型可分類為「監督式學習」(Supervised Learning)、「非監督式學習」(Unsupervised Learning)、「半監督式學習」(Semi-supervised Learning)、「增強學習」(Reinforcement Learning)等。機器學習模型演算法的數量與方法非常多，包括了類神經網路(Artificial Neural Network, ANN)、支援向量機(Support Vector Machine, SVM)、集群(Clustering)、決策樹(Decision Tree)、隨機森林(Random Forest)….等等。

面對層出不窮、利用新技術的資安威脅，越來越多的資安廠商也開始將機器學習技術應用在強化網路安全防護。科技可以為善、也可以作惡，端看人類如何使用它，人工智慧、機器學習技術也是如此。在資安世界裡，這場防禦端和攻擊端的人工智慧攻防前哨戰已然開打。

機器學習在DDoS安全防禦技術的應用

一直以來，威睿科技的DDoS防禦技術，不同於傳統特徵碼比對的資安防禦技術，是利用全網大數據流量的採集，進行大規模的流量行為分析(Network Behaviour Analysis)佐以流量基線動態學習(Automatic Baseline Learning)，來進行全網的DDoS檢測及防禦。近年來大數據(Big data)及機器學習等IT技術逐漸成熟，威睿科技也積極投入運用機器學習提升現有DDoS攻擊檢測技巧的相關研究。我們主要的研究方向有：

高效流量資料收集，生成流量屬性特徵(Feature)

每個用戶的網路行為，實際上可以被拆解為一條條的流量記錄(Flow record)，透過底層的網路設備傳送出來。GenieATM系統能從骨幹網路路由器採集到流量摘要記錄，每則記錄包含IP地址、協議、端口、TCP標籤、大小、時間等數十個欄位，是一種基本的數據來源；基於這些採集到的資訊欄位，我們可以將資訊加以聚合、分類、進行關聯(correlation)，例如把各條流量的各種屬性，關聯至其他資訊如國別、子網、BGP路由屬性等等；也可以做出預處理如進行分類、統計及排序等，以得出更多屬性特徵的數據；另外，GenieATM系統既有的DDoS事件記錄(anomaly ticket)則是另一種數據來源，亦可以由此關聯得出更多的屬性特徵。

在觀測一個流量突發的特徵時，一般會使用流量的累計值作為指標，這是在運營商等級的數據採集偵測上常用的指標。針對流量其他特徵的排序分析，往往被用在流量分析上、而不是異常偵測上。然而在機器學習的應用上，我們會試著將原本是一組排序數列的分析結果，轉化成一個單一指標數字。這個指標數字可以簡化表示在某一時間序列(Time series)的一種狀態，如國別分佈狀態、應用分佈狀態、運營商來源分佈狀態.....等等，用以比較在不同時序間(每日、每週、每週末.....等等)，該種狀態的差異性。當某種狀態的差異性被檢測到時序狀態上有「異常」時，迅速提出異常流量行為的告警。

透過這種總合式流量數據及既有攻擊記錄關聯來準備機器學習的訓練資料(Training data preparation)，讓機器學習看的不只是單線的行為，而是全面的趨勢特徵與方式。

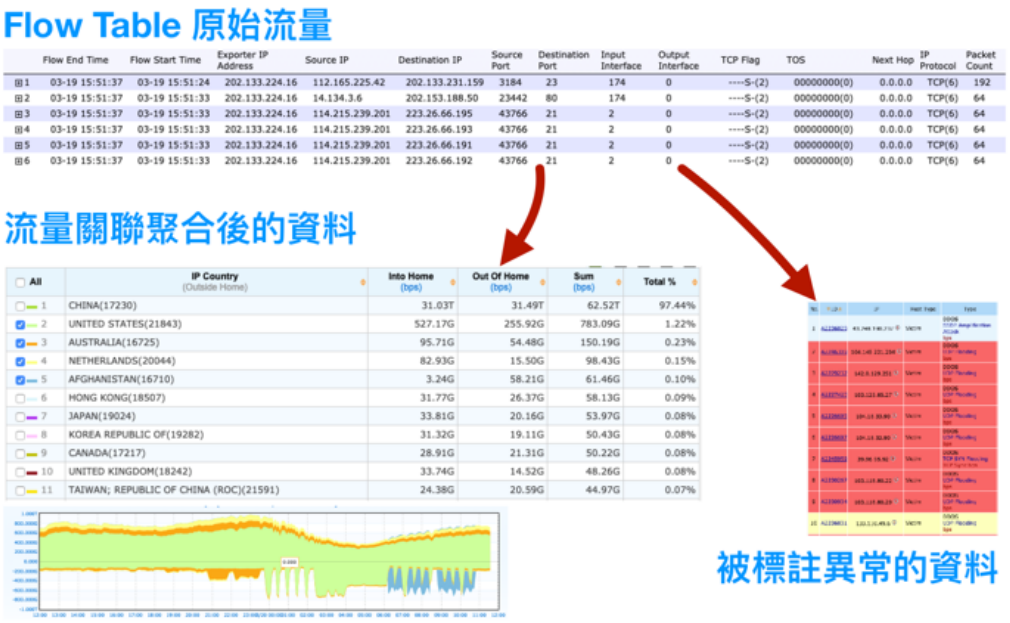


Figure 1: 流量屬性特徵處理(Feature Engineering)

持續調校機器學習模型，不斷優化機器學習結果

在異常流量檢測的機器學習應用上，就是要透過特徵工程(Feature Engineering)去將適合的流量屬性特徵放入適合的機器學習演算法進行分類及分析。當我們的訓練資料集裡有很多個特徵時，首先要萃取出最具影響的幾個特徵來做分類器模型的訓練，而不要使用所有可能的流量屬性特徵數據來做訓練，也就是所謂的進行「降維(dimension reduction」。這時我們可以利用成分分析(Principal component analysis，PCA)來萃取特徵。主成分分析(PCA)是一種特徵提取的演算法，利用特徵降維來避免因維度災難(curse of dimensionality)所造成的過度適合(Overfitting)現象。

此外，網路流量的資料，本身就是一種時序性(Time-series)的資料。所以，在分析以進行異常檢測或預測時，就不得不考慮時序性的學習和分析模型。比方說，最簡單的流量基線模型，就是將即時的流量行為，去與過去持定時間區段的流量行為做統計性的比對，如現在和過去一週、一個月的流量。然而，這特定時間區段的選定，通常是人為選定的、固定的、不具自動智能的。但在機器學習技術的幫助之下，我們得以考慮時序性資料的更多組成因素，如基線閾值(baseline level)、線性趨勢(linear trend)、規律性(sesonality)和噪音(noise)因素等。常被用於時序性資料分析預測的機器學習模型有自動迴歸整合移動平均(Auti-Regressive Integrated Moving Average, ARIMA)、長短期記憶網路(Long Short Term Memory Network, LSTMN)等等。

一般機器學習運用於DDoS異常流量檢測時，在沒有標籤資料(labeled data)的情形下可以使用無監督式的學習，將流量以萃取出來的流量特徵進行自動分群，以找出「離群者(outlier)」，離群者指的是那些不符合其他資料特徵行為模式的資料點，也就是我們要檢測出的異常流量。這類常用的機器學習演算法，有K-means集群分析(K-means Clustering)、孤立森林異常檢測 (Isolation Forest)、局部異常因子演算法

(Local Outlier Factor)等等。

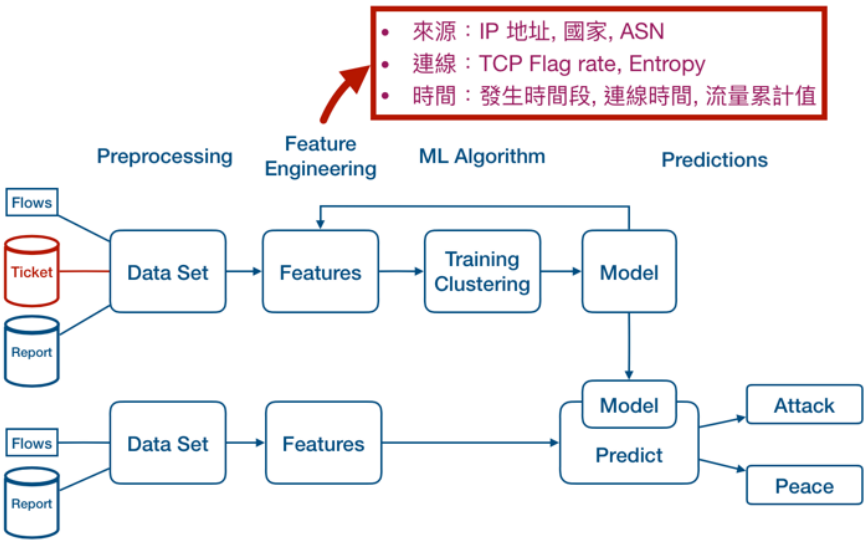


Figure 2: 非監督式(Unsupervised)學習程序示意

做為一個持續檢測DDoS攻擊的即時系統，也能同時利用已有方式測得的攻擊事件作為回饋輸入，做成多層的監督式學習模型。例如將既有攻擊事件的流量屬性特徵，自動回饋至機器學習的判定標籤上，讓機器學習的準確度提升，同時免除人工操作標籤資料的麻煩及缺乏效率。

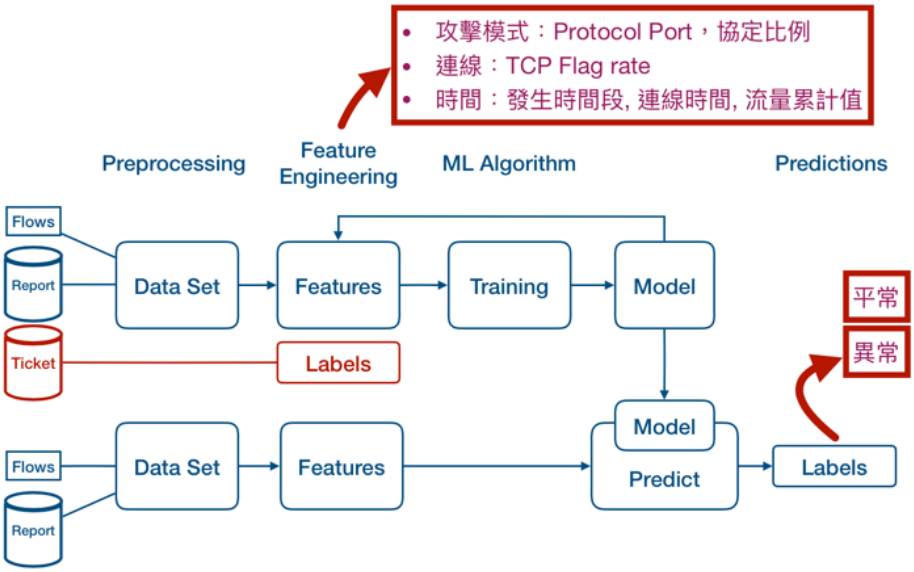


Figure 3: 監督式(Supervised)學習程序示意

視覺化機器學習歷程結果

將流量屬性與趨勢，透過視覺化的效果展現，讓資料在處理後有效地將資料快速地呈現在眼前，讓使用者能簡單明瞭分析後的資訊所帶來的商業價值。同時也將機器學習的歷程與結果，透過圖形與動畫，提供互動式的操作讓使用者以直覺方式修正模型參數，免除對於演算法，數學算式，與統計數據的疏離感。

Figure 4: 網路威脅視覺化儀表板

結語：以機器學習打造新一代DDoS安全防護系統

近年資安威脅情勢持續升高，驅使資安解決方案積極廠商強化其防護能力。過去的的資安技術在許多面向上仰賴所謂的領域專家，如產出攻擊行為特徵碼、如相關系統閾值的設置與調校等。然而，人類專家難以擴展。你可能希望能夠自動將某種流量行為與其他流量行為進行比較，尤其是當所謂的流量行為可以有許多種不同特徵時，不論是在流量內容上的、流量位置上的、流量速率上的、或是時間序列上的。人類專家無法以人力分析和比較所有的流量行為特徵，並且當流量規模大、網路架構複雜、流量行為多樣化時，就更不可能仰賴人類專家了。此時，資安廠商也開始引進機器學習技術的應用。

由於基於機器學習的人工智慧技術能夠分析更複雜、非結構化的資料並解決難以定義規則的問題，可以協助處理IoT的各種聯網載具、多樣化的網路應用、大量的異質數據，及更重要的是越來越多的自動化智慧型資安攻擊行為的複雜狀況，因此使用機器學習的資安防護技術已然成為相關解決方案業者的發展重點。

威睿科技長年耕耘於電信網路運營商市場，在電信級IP網路流量分析與DDoS防禦技術領域累積多年的經驗。在現今大數據(Big data)及機器學習(Machine Learning)等IT技術的不斷創新及日漸成熟的技術支撐下，將這些新技術融合、應用於威睿既有的解決方案中，是必將撰擇的道路。透過大量的網路流量資料收集、從中萃取屬性特徵訓練資料、選取適合的機器學習演算法，讓機器自動學習建立多向量的攻擊檢測模型，據以進行攻擊事件的即時檢測，最終目標是希望讓系統能像人類專家般，防禦系統也可自動藉由事件處理的經驗累積達到學習成長，以追求最快速又準確的DDoS安全防護，協助客戶打贏這場越趨嚴峻之資安AI攻防戰。

GenieNetworks

威睿科技股份有限公司 | 地址：台北市內湖區內湖路一段360巷15號5樓
電話：+886 2 2657 1088 | 官網：www.genie-networks.com/?lang=ch