<Newgras Blog>

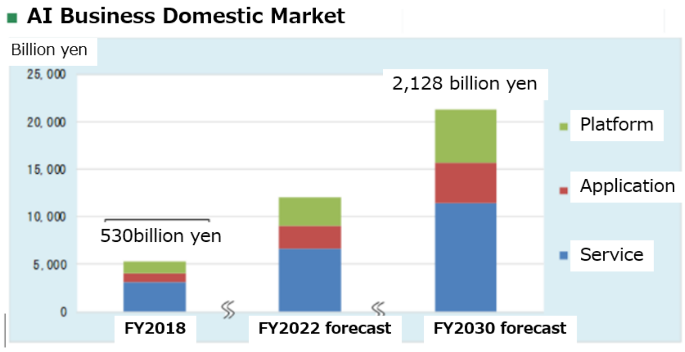
Significantly improved user operation management by applying AI/ML algorithm

～Causal Dependency & event inferencing adaptive monitoring～

Evangelist Hitoshi Yoshioka

Introduction

Since several years ago, AI (artificial Intelligence) has been introduced in a wide range of fields, and the AI business market is expanding. According to a study by Fuji Chimera Research Institute, AI-based analysis services, AI-based consulting, and AI-incorporated applications and edge computing platforms have been developed. According to this Institute's "2019 Artificial Intelligence Business Survey", the market forecast for FY 2030 is expected to be ＄21.2 billion for AI business. However, the level of AI varies widely. AI-like machine learning (ML) is also increasing. The author thinks that this is merely pattern matching that statistically analyzes various pattern data and automatically searches as needed, and that it is not the original use of machine learning-based AI. AI/ML-based products are also increasing in the network field. Up to now, the threshold of traffic (“baseline”) is not just for network data but for any performance data. In the emerging models, the traffic data flowing on the network and other performance data has to be learned, and then the threshold of the normal state is automatically set, i.e., normalize unnecessary alerts. It can be said that this is a function to which AI/ML is applied. However, the author does not want to call this level of automation as AI/ML. It is just a learning function. Among them, the author would like to explain the advanced AI/ML recommended by the author based on actual cases.



Source: Fuji Chimera Research Institute “2019 Artificial Intelligence Business Survey”

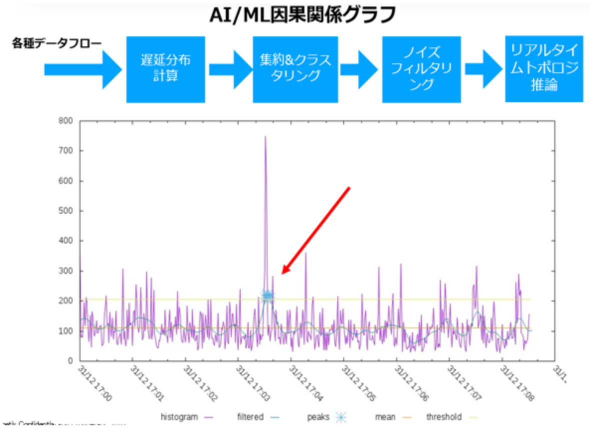
Ennetix AI/ML algorithm overview

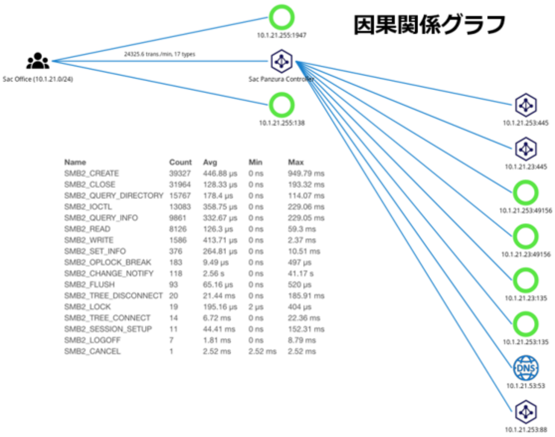
Ennetix is a university-initiated venture company developed and launched by a computer science professor at UC Davis and his group. The detailed profile of Ennetix founder Dr. Bis Mukherjee is omitted here, but he is still a professor at the university and has been a world leader in optical and broadband access networks for over 30 years.

The AI/ML algorithm of Ennetix is composed of the following components.

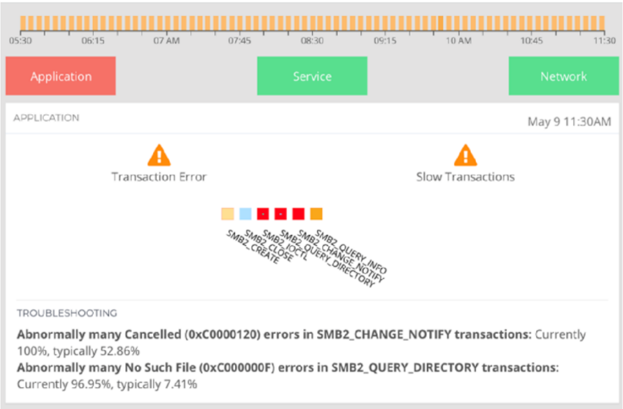
■Causal Dependency Graph

It is used for automatic detection of components (server, router, DNS, LDAP, etc.) that have a Causal Dependency with the application access infrastructure for each user group. Causal Dependency indicates the relationship between the “cause” and the “effect” of how one element affects another one.





Name	Count	Avg	Min	Max
SMB2_CREATE	39327	446.88 μs	0 ns	949.79 ms
SMB2_CLOSE	31964	128.33 μs	0 ns	193.32 ms
SMB2_QUERY_DIRECTORY	15767	178.4 μs	0 ns	114.07 ms
SMB2_IOCTL	13083	358.75 μs	0 ns	229.06 ms
SMB2_QUERY_INFO	9861	332.67 μs	0 ns	229.05 ms
SMB2_READ	8126	126.3 μs	0 ns	59.3 ms
SMB2_WRITE	1586	413.71 μs	0 ns	2.37 ms
SMB2_SET_INFO	376	264.81 μs	0 ns	10.51 ms
SMB2_OPLOCK_BREAK	183	9.49 μs	0 ns	497 μs
SMB2_CHANGE_NOTIFY	118	2.56 s	0 ns	41.17 s
SMB2_FLUSH	93	65.16 μs	0 ns	520 μs
SMB2_TREE_DISCONNECT	20	21.44 ms	0 ns	185.91 ms
SMB2_LOCK	19	195.16 μs	2 μs	404 μs
SMB2_TREE_CONNECT	14	6.72 ms	0 ns	22.36 ms
SMB2_SESSION_SETUP	11	44.41 ms	0 ns	152.31 ms
SMB2_LOGOFF	7	1.81 ms	0 ns	8.79 ms
SMB2_CANCEL	1	2.52 ms	2.52 ms	2.52 ms



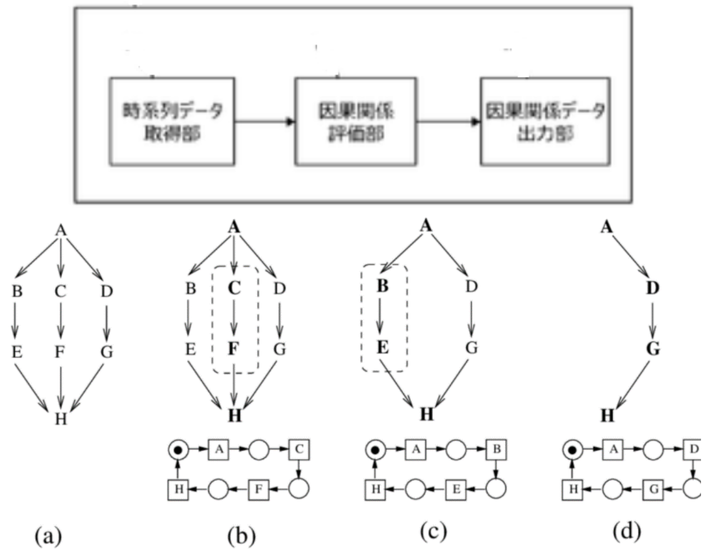
TROUBLESHOOTING

Abnormally many Cancelled (0xC0000120) errors in SMB2_CHANGE_NOTIFY transactions: Currently 100%, typically 52.86%

Abnormally many No Such File (0xC000000F) errors in SMB2_QUERY_DIRECTORY transactions: Currently 96.95%, typically 7.41%

[ここに入力]

A causal dependency is an event in which B changes due to A. This is a method of expressing the causal relationship between multiple events (elements) (called a causal network) in a table or graph and using this for behavior prediction or cause estimation when a problem occurs.



Note: What is the difference between correlation and Causal Dependency?

Correlation means that something between A and B has something to do with something, and Causal Dependency means that B fluctuates because of A as a "causal relationship." That is, the causal relationship is a three-dimensional and variable perspective, whereas the correlation is a planar perspective. In the service that provides end-to-end quality of experience for users, the automatic detection function based on the causal relationship by AI/ML is important for the failure of the causal relationship with the important application access infrastructure of each user group.

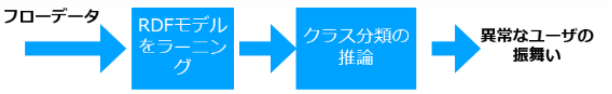
For example, the relationship between two quantities is shown in which when one increases, the other tends to increase or decrease. In the correlation that is often used for monitoring in APM (Application Performance Management) / NPM (Network Performance management), by comparing the predictions from the model with the current situation, we find "unusual" behavior and make the association when assisting identification. In short, it is the detection of correlation destruction. NPM uses the Root-Cause Analysis (RCA) function to investigate the correlation of a large number of events, and analyze the cause of failure based on the Layer 2 and Layer 3 topologies. On the other hand, in the correlation of APM, information on the Web, application server, and database is matched to visualize the overall application usage status from transaction information and identify the cause. However, in end-to-end application access via a network, a complicated combination of network, network services (DNS, LDAP), application (SAP, O-365, etc.) exists in actual network access. To include this, it is necessary to combine NPM, APM, and *network service tools in a complex manner.*

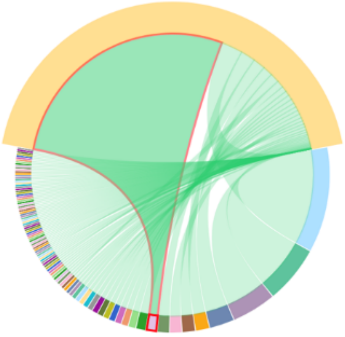
■Packet-Train Technology

Packet-train technology is a performance method in which multiple probe packets are continuously transmitted to the network, and the bandwidth is directly measured by transmission delay, traffic loss, and packet interval on the bottleneck link. Specifically, the traffic stream to be measured is packet-captured by a mirror port connection to estimate the end-to-end route and the bandwidth, capacity, and hop-by-hop network characteristics.

■Random Decision Forest Models

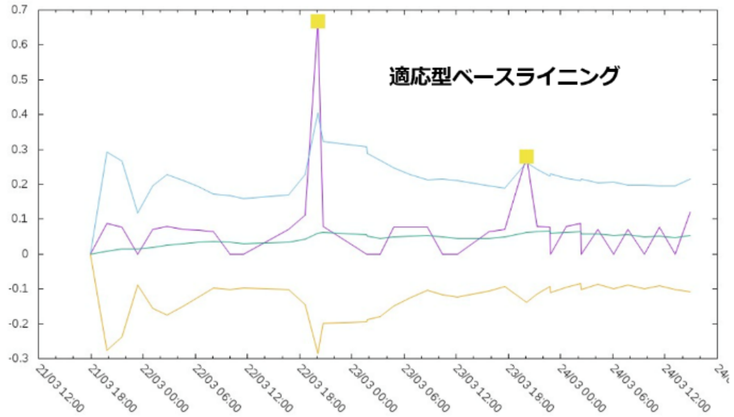
In random forest, a large number of decision tree models are created by random sampling that allows duplication, and the final prediction value is determined by majority voting of the prediction results of each tree. Specifically, it determines the user traffic pattern and characteristics called "behavior".





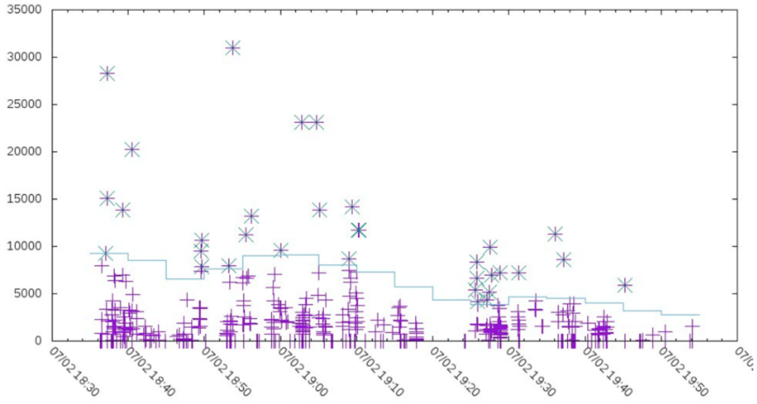
■Parametric Regression

Parametric regression obtains a model (mathematical formula) that represents data well by assuming a model to some extent and estimating the model parameters that best fit the given data. Specifically, automated adaptive baseline and user behavior events and patterns are identified and alerts are generated by this algorithm.



Ennetix products are targeted for mission-critical applications and apply true AI/ML analysis algorithms, focusing on user experience (“user experience quality”). There are various levels of AI/ML user experience analysis algorithms. Detailed technical explanation is omitted here, but the unique AI/ML algorithm applied in Ennetix is explained with reference to a case example:

In this use case, 20 kinds of user applications are used. Here, abnormalities in performance and security that have different characteristics (e.g., number of users, alerts, throughput, etc.) for each application are automatically detected and diagnosed. And with Ennetix products, the alert analysis man-hours of the operation manager are greatly reduced. Then, using the advanced AI/ML causal dependency function, the performance hotspot analysis is accurately specified. The figure shown below is an example of clustering analysis targeting large-scale causal dependency graph data, and AI/ML analysis such as classification is performed only on measurement data.



In this way, we perform inference analysis of end-to-end real-time topology applying AI/ML. It intelligently infers performance and security deviations, and identifies the root cause in real time. The best-practice information provided by the Ennetix AIOps solution helps to enable the user experience for mission-critical applications.

Note: *Clustering analysis is a method of classifying and grouping, from a group in which different properties are mixed, those with similar properties according to certain rules and common terms.*

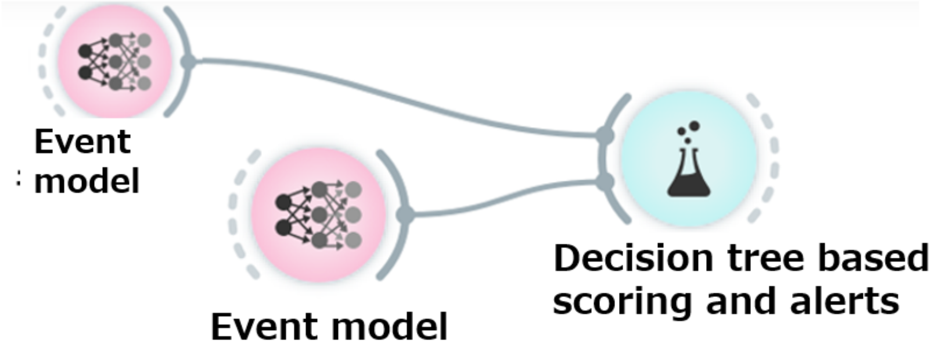
In this use case, some applications generate 20-30 alerts per month, while others have high volume alerts. Also, all alerts are analyzed by applying AI/ML algorithms. Generally, in APM/NPM, a threshold value (baseline) is set manually or semi-automatically to generate an alert. Thousands of huge alerts will result if no threshold is set. In this use case, you can see how many users are using which applications, and even this simple information is valuable in this use case. And when a failure occurs, we want to quickly identify the content and cause of the failure (whether it is an application, a network, or a network service).

The Ennetix use case doesn't just stick to alerts. It is very useful when there are failures in network services such as DNS and LDAP. Therefore, we consider these alerts as service layer alerts. This is also a unique feature and definitely a "necessary and sufficient" function.

In addition, AI/ML models are used to identify abnormal traffic and user application behavior.

■Alert generation and a method to specify alert

Ennetix continuously generates thresholds for all the collected performance metrics. These performance metrics include delay, loss, jitter, bandwidth, retransmissions, and so on. If the current data point of the performance metric is higher than the threshold, the Ennetix solution creates what is called a performance event. There is another AI/ML algorithm that captures these performance events together with parameters into one model, and decides whether to generate an alert (notice, warning, or critical) based on parameters, time period, and other characteristics.



Conclusion

The application of AI (artificial intelligence) was explained at the beginning, but there are machine learning and deep learning, and as its name implies, machine learning models regularity in a large amount of data, and until now people have been good at letting machines do the work. Deep learning is a technology that goes one step further than machine learning, and is designed to allow the machine to have more difficult recognition, cognitive, and discriminative functions, called a neural network, which is designed to simulate the behavior of the nervous system of living organisms. The deep learning is executed by an object with an order of magnitude higher than that of the machine learning, and can be used to accurately identify image recognition or security behavior. The author expects that Ennetix's next-generation application-type RCA (Root-Cause Analysis) applying AI/ML algorithms will greatly contribute to countermeasures against the ever-increasing complexity of user-specific application access and improve user experience quality.