June 2021

# DISTRIBUTED PARAMETER MONITORING FOR WIRELESS DEPLOYMENTS

Mirko Raca

Domenico Ficara

# DISTRIBUTED PARAMETER MONITORING FOR WIRELESS DEPLOYMENTS

AUTHORS:
Mirko Raca
Domenico Ficara

## ABSTRACT

Using machine learning (ML) to make observations of network operations is faced with many constraints, including collection constraints, storage constraints, and processing constraints. Additionally, in many instances, data collected from a network may be unusable and will incur collection, storage, and processing costs with potentially limited return. Presented herein are techniques through which pre-filtering tasks can be distributed to wireless access points (APs) to highlight valuable metrics and learn from network deployments.

## DETAILED DESCRIPTION

Although devices deployed in a network can provide plenty of data, the reality is that only one tiny portion of this data may be useful for formulating network analytics and extracting actionable business value out of such analytics. For example, ML can be used to make observations regarding network operations, but is faced with various constraints, including:

- Collection – a lot of data is often needed to perform network analytics, which can be difficult to detect due to network bandwidth issues;
- Storage – data gathered from a network is to be exported and stored in a central repository in order to perform network analytics; and
- Processing – considerable computing resources are often involved for performing network analytics, which is why such analytics are often exported to the cloud (potentially causing privacy issues).

Further, it is a hard reality to face that it is possible that up to 99% of data collected from a network may be unusable; yet, will incur collection/storage/processing costs with potentially limited return. Such issues pose challenges both on an analytics engine to extract a useful feature-set and also on a network to stream telemetry data from network devices to the cloud or on-premise analytics devices/engines.

This proposal provides techniques through which wireless access points (APs) and wireless local area network (LAN) controllers (WLCs) can be used as edge-node computing resources to preprocess metrics and "float-up" meaningful parameters/data in order to minimize bottlenecks and/or speed-up decisions/conclusions of network analysis. In essence, the techniques presented herein provide for the ability to share the load of network analysis and distribute the evaluation and extraction of valuable metrics with network devices, such APs and WLCs.

Consider a set of interactions, as illustrated via the example system of Figure 1, through which various techniques of this proposal may be illustrated with the goal of understanding the effects of an unknown number of features on a goal metric.

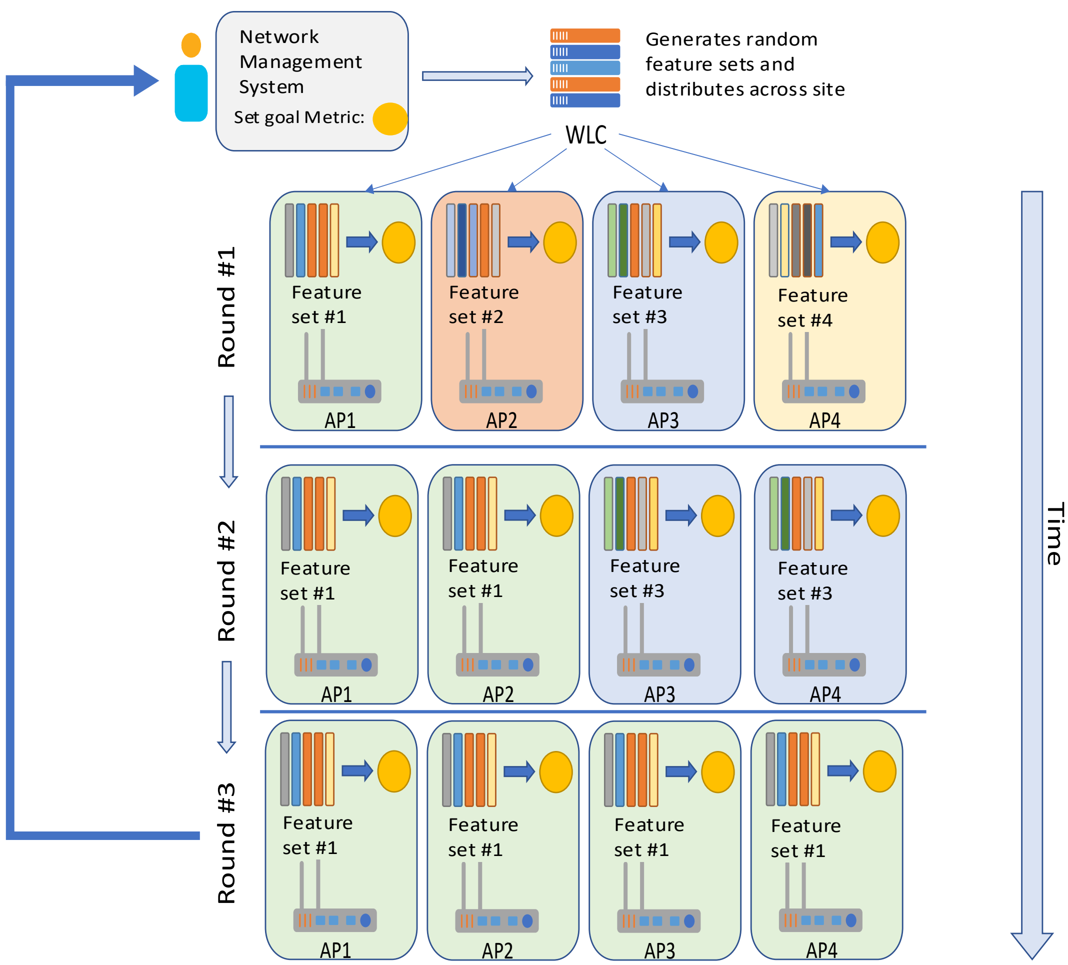


*Figure 1: Example System Operations*

As illustrated in Figure 1, consider that a network administrator wants to measure the effects on a chosen goal metric (e.g. on-boarding time of clients, client ping rate, etc.), and assigns the system to monitor the goal metric through a user interface (UI) of a network management system. Next, the system distributes the tasks of monitoring the goal metric to all edge-nodes (APs) and assigns to each AP a random subset of system metrics (features) to analyze. At this step, each AP acts on a single unique hypothesis. In this example, consider one hypothesis "Hypothesis25," which assumes that "metrics (x,y,z,t) affect the on-boarding time of client."

Thereafter, the APs collect evidence (e.g., (x,y,z,t -> onboarding_time) pairs) to support the connection between the goal metric and set of the hypothesis. The evidence might be in the form of statistical tests of correlation between the features and the goal metric (e.g., t-tests, chi-squared tests, etc.) or regression fits (e.g., linear correlation between 2 values).

The APs that do not find a significant correlation in a predefined duration, which can be expressed as the number of connections, clients or hours, are assigned a new hypothesis (again random feature combinations) for testing for the duration of the experiment. If an AP determines a significant correlation, it can request from the network management system to perform a broader sample. The number of APs testing the Hypothesis25, thus, increases from 1 to 25, for example, and the test continues with more evidence. If there is significantly broad support (e.g., enough client events/devices supporting Hypothesis25) and correlation is significant, the conclusions can be presented to the network administrator.

Regarding workload distribution, the system metrics can be locally measureable on each AP in one instance. This means that each AP works only on data from its own clients and its own information (e.g., central processing unit (CPU) information, memory, wireless metrics such as Received Signal Strength Indicator (RSSI) / Signal to Noise Ratio (SNR) / etc., geographical location, and/or the like).

In another variation, the metrics can be collected from close-by APs and shared cooperatively. In this case, one AP can be elected as the AP that will perform the data-processing and participate in an experiment. Such an AP can be chosen (i) randomly, (ii) according to its CPU/memory capability, (iii) according to its current workload. In some instances, different elected APs can be chosen in the same location for different experiments. In another variation, the processing of multiple experiments can be performed at different elected APs at the same time (when CPU/memory and workload make it possible) and each AP can share its set of metrics with elected APs according to the experiment for which they are responsible.

The system can present/report conclusions in several forms. In one instance the system can illustrate the major effects on a given goal metric to a network administrator that take various actions. For example, if onboarding time is significantly affected by the number of clients connected to an AP or if on-boarding is affected by the throughput at the moment of on-boarding, an administrator can take various actions, such as, respectively, increasing the number of APs or limiting the bandwidth of connected clients. In another instance, the system can store a query for permanent monitoring. For example, if a meaningful metric is presented to a network administrator that cannot be immediately changed/affected, the administrator may choose to monitor the metric over time for future events.

Various advantages and/or benefits may be realized in accordance with the techniques described herein. For example, the techniques described herein may facilitate distributed data processing such that data does not have to be transferred to a centralized location for processing by a centralized processing resource, rather, the processing power distributed edge nodes, such as wireless APs, can be utilized to determine system metrics. Additionally, techniques herein are not limited to observations for a predefined set of "significant features," rather; larger sets may be utilized using random feature combinations. Further, the process described herein can operate automated and unsupervised, having minimal need for retraining network administrators.

In summary, techniques herein may involve the random selection of system metrics/features that can be distributed to edge nodes or APs for evaluation. Further, experiments can be distributed across APs in different manners that allow the APs to work with locally available and/or close-by data through different, flexible data analysis scenarios. Finally, automatic adaptation is provided that evolves and highlights the best metrics in order to learn from and optimize network deployments.