AutoFocus: Automatically Scoping the Impact of Anomalous Service Events

Ren Quinn
University of Utah
renquinn@cs.utah.edu

Zihui Ge
AT&T Labs - Research
gezihui@research.att.com

He Yan
AT&T Labs - Research
yanhe@research.att.com

Jacobus Van der Merwe
University of Utah
kobus@cs.utah.edu

I. Introduction

Networks, and the services they enable, are increasingly diverse and highly utilized. From DSL and fiber-to-the-home access networks, to cellular mobile networks, to content-delivery networks; all require extensive monitoring in order to meet the increase of user expectations of the availability and quality of those services provided to them. The complexity of these networks and services require better management on the part of providers as the data resulting from service monitoring experiences an increase in dimensionality, making it difficult to fully interpret anomalies in the data. For example, anomaly detection generally says "I found an anomaly with mobile phone A in market Z". But it is more useful to know what other phones and what other markets are also experiencing the same anomaly.

We call this problem **impact scoping** of anomalous service events. It is the first important troubleshooting step when operators receive notification of an anomalous event, at which time, the full scope of the impact of the event is often unclear. This troubleshooting process occurs across various network services. E.g., when analyzing problems using customer data from a *video transmission service*, one might ask questions about the set top box the customer uses, the region in which they reside, if a specific channel(s) is/are experiencing problems, or what type of service plan the customer subscribes to. Similarly, when problems arise in *content delivery networks* (CDNs), operators might ask questions about what type of content is being accessed, what areas are affected, or what client devices/software is affected.

With the enormous size of these datasets plus their increasing dimensionality, determining the impact scope of anomalies is becoming increasingly complex. Because the number of domains in which this process is used, and the similar nature of troubleshooting questions asked in multidimensional scenarios, we see the need for a generic approach to automate the process for all domains. In this paper we present such an approach, called AutoFocus: an algorithm which, given a single instance of an anomalous event, looks for symptoms of that anomaly across multiple dimensions (region/market, phone model, etc.). The key is our *fitness score* metric which explains the characteristics of an anomaly. It represents the level of anomalous behavior for a given time series with respect to other behaviors across dimensions in the dataset.

Finding the scope of an event then becomes an exercise of finding an aggregate time series which has the highest fitness score (aggregate meaning a combination of time series corresponding to the attributes, or values, of a dimension). However, given the size of typical monitoring datasets, testing every aggregate combination for the highest fitness score is prohibitively expensive. We developed AutoFocus to streamline this process by significantly trimming the search space associated with an event. Specifically, AutoFocus searches for time series aggregates that produce the locally maximum fitness score, then iteratively identifies attributes from each dimension that increase the overall fitness score.

The contributions of this paper are as follows:

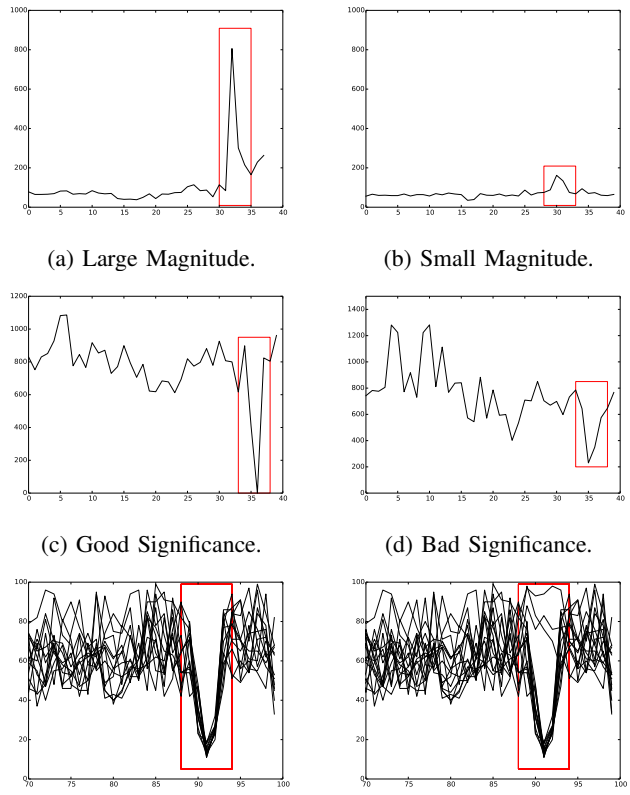
- A fitness score metric, which models how well an anomaly fits across relevant dimensions of a multidimensional dataset.
- An algorithm which, guided by the fitness score, significantly reduces the search space in the process of impact scoping.
- We show that our approach provides a general solution to the problem of scoping the impact of anomalies.
- We evaluate our approach using three diverse, real datasets.

II. Methodology

Given an instance of an anomalous event, the goal of impact scoping is to understand the scope of the impact of that event. We developed a fitness score metric to help automate this process by interpreting the behavior of an anomalous event buried within a multidimensional, time series dataset. It analyzes the **attributes** of the anomaly (i.e., the possible values or instances of a dimension, such as Utah in the Markets dimension), and then can be used to search for similar behaviors across all attributes for all dimensions in the dataset.

The idea of a fitness score comes from genetic algorithms where the evolutionary process of natural selection or "survival of the fittest" is used to find the solution to problems with many possible solutions [2], [11]. They trim the search space to only potential solutions with a higher probability of being the correct solution, all driven by a fitness score which determines the strength, or *fitness*, of a particular member of a population. Or in the case of search algorithms, how close a potential solution is to the optimal solution. Our fitness score serves the same purpose: we wish to find the subset of our data that is more relevant (i.e., higher fitness) to a given anomaly.

978-3-901882-98-2



(a) Large Magnitude. (b) Small Magnitude.

(c) Good Significance. (d) Bad Significance.

(e) Good Consistency. (f) Bad Consistency.

Fig. 1: The fitness score filters out time series which violate each statistical component. These graphs, drawn from our Cloud API dataset, illustrate each component.

We define "relevance to an anomaly" in terms of four main components which explain different characteristics of an anomalous event (illustrated in figure 1):

Magnitude. Determines how much of the event is captured by a subset of the dataset, and if the subset encompasses all elements pertaining to the event. Larger magnitudes suggest higher fitness scores. Figure 1a shows an obvious and significant spike in the time series. Whereas the spike in figure 1b is not very prominent, it could even be just noise.

Statistical Significance. Determines how significant the event is as compared to normal variations. Figure 1c shows a time series with an event of significant magnitude with respect to the historical variance. However, the event in Figure 1d is less meaningful due to the high historical variance in the series.

Consistency. Reveals the level of difference between components of the scope. Figure 1f is a collection of time series to be analyzed for anomalies. However, it includes a few time series that do not follow the same anomalous patterns as the rest. This aggregate has low consistent as it does not accurately represent the anomaly. Whereas, in figure 1e, with the irrelevant time series removed, there is little difference between the components resulting in a higher consistency.

Information Criteria. Defines the simplest explanation of the scope of the event, the parameters. In other words, it says how much of the scope is included in the dataset.

A. Notation and Fitness Score

Consider a networked service and its service performance metrics defined over multidimensional attributes. We define such a service performance metric as $\mathbf{f} : \mathbf{A_1} \times \mathbf{A_2} \times \ldots \times \mathbf{A_k} \rightarrow \mathbf{R} \cup \{null\}$ where for $i \in [1, k]$, $\mathbf{A_i} = \mathcal{P}(\mathbf{a_i})$ represents the superset of $\mathbf{a_i}$, which is the set of values for the i-th dimension attribute. For example, the end-user phone model attribute may take values in $\mathbf{a_i} = \{$pc, tablet, smartphone, misc$\}$. For example, let $\mathbf{f}$ be the average page loading time, and $\mathbf{f}(\{$App A$\}, \{$Smartphone$\}, \{$Phone Model A, Phone Model B$\})$ reports the average page loading time for the service requests originated from App A running on either Phone Model A or Phone Model B. When no service matching the specified attribute values are observed, $null$ value is expected. An example of such incompatible attributes is $\mathbf{f}(\{$App B$\}, \{$tablet$\}, \{$Phone Model A, Phone Model B$\}) = null$.

While $\mathbf{f}$ is time varying, typically modeled as a time series, we simplify the time series aspect of $\mathbf{f}$ in our application as follows. We define the observed performance metric as $\mathbf{f}_{t_s,t_e}$ during the detected anomalous service event time (from t_s to t_e), and define the expected performance as $\widehat{\mathbf{f}_{t_s,t_e}}$. There exists a rich literature on multidimensional time series forecasting [7], [6], [10], [8] that can be leveraged to derive $\widehat{\mathbf{f}_{t_s,t_e}}$ from historical values of $\mathbf{f}$. For simplicity of the presentation, we drop the time t_s, t_e annotation such that both $\mathbf{f}$ and $\widehat{\mathbf{f}}$ refer to the service performance during the service anomaly.

The scoping of the service problem is to identify x where $x \in \mathbf{A_1} \times \mathbf{A_2} \times \ldots \times \mathbf{A_k}$ such that all services matching x are affected by the service anomaly while all services out of the scope $(\mathbf{a_1} \times \mathbf{a_2} \times \ldots \times \mathbf{a_k} - x)$ are unaffected.

Since it is challenging to determine whether a service is actually affected by the service anomaly event without the knowledge of the root cause, we construct a set of heuristics that guide us in finding the right scoping x.

The service degradation *magnitude* function H_m:

$$H_m(x) = |\mathbf{f}(x) - \widehat{\mathbf{f}}(x)|$$

captures the degree of service degradation in the scope.

The *significance* of service degradation function H_z :

$$H_z(x) = \min\left(\frac{|\mathbf{f}(x) - \widehat{\mathbf{f}}(x)|}{\delta(x)}, z_{max}\right)$$

captures the significance of the observed service degradation compared to the historical performance, where $\delta(x)$ is the standard deviation of the forecast error of $\widehat{\mathbf{f}}(x)$, and z_{max} is used to guard against small or zero $\delta(x)$ causing unbounded $H_z(x)$. This is essentially the z-score of the service degradation. We use $z_{max} = 5$ for the remainder of the paper.

The *consistency* of service degradation function H_c:

$$H_c(x) = \sum_{i=1}^{k} H_c^{(i)}(x)$$

captures the degree of variations in service degradation among the sub-scopes of x, where k is the number of attribute dimensions and

$$H_c^{(i)}(x) = 1 - \sum_{y \subseteq_x^{(i)} x} \left| \frac{|\mathbf{f}(y) - \widehat{\mathbf{f}}(y)|}{\sum_{z \subseteq_x^{(i)} x} |\mathbf{f}(z) - \widehat{\mathbf{f}}(z)|} - \frac{\mathbf{f}(y)}{\sum_{z \subseteq_x^{(i)} x} \mathbf{f}(z)} \right|$$

Here $y \overset{(i)}{\subseteq} x$ denotes that y differs from x only in the i-th dimension and the value of y in the i-th dimension has cardinality of one (i.e., has a single element) and is a subset of the value of x in the i-th dimension. For example, both $\{a,b,c\} \times \{d\} \overset{(2)}{\subseteq} \{a,b,c\} \times \{d,e\}$ and $\{a,b,c\} \times \{e\} \overset{(2)}{\subseteq} \{a,b,c\} \times \{d,e\}$ are valid.

And finally the *information criteria* function H_i:

$$H_i(x) = \sum_{i=1}^{k} C - \frac{|x^{(i)}|}{|\mathbf{a_i}|}\left(1 - \frac{|x^{(i)}|}{|\mathbf{a_i}|}\right)$$

provides a bias toward simple scoping construction, where $x^{(i)}$ is the value at the i-th dimension of x and $C > 0.25$ is a weighting constant. Note H_i is maximized if all or none of the components in every dimensions are affected. This heuristic observes the Occam's razor principle, that is, the simplest answer is most often correct.

The overall fitness score of scoping x is a combination of the four heuristics above: $H(x) = g(H_m(x), H_z(x), H_c(x), H_i(x))$ It is not hard to construct scenarios in which the function g should be best tailored to give weighting to different considerations. For simplicity, we use the product of the four heuristic functions as the fitness score of any given impact scoping x.

B. AutoFocus

We designed AutoFocus to automate the process of quickly finding the scope of an anomaly using our fitness score. Given the date and time of a reported anomalous event as input, the main idea is to find the entire scope of the event by checking the fitness score of different combinations of time series aggregates. The aggregate with the maximum fitness score then represents the entire scope of the anomaly.

AutoFocus filters through the relevant attributes of each dimension until it finds the aggregate time series which represents the correct scope of the given event. Starting with the dimension with the single attribute with the highest fitness score, we select only relevant attributes from that dimension. To do this we combine all attributes of the dimension, one at a time (starting with the attribute with the highest fitness score), until an attribute causes a decrease in the running fitness score.

Having identified the most relevant attributes of a single dimension, AutoFocus continues to search the remaining dimensions for relevant attributes in order of which dimensions have the highest individual attribute. However the difference in the following iterations (i.e., dimensions) is that the attributes previously identified as relevant are maintained as context under which any future fitness score calculations are performed. This is key to trimming down the search space for all relevant attributes in all dimensions. It makes sense intuitively as well; once an attribute is deemed irrelevant to an event, it should be omitted when deciding the relevance of other attributes.

The last step within each iteration is to reorder all remaining dimensions, again by maximum individual attribute fitness. Adding the context of fixed attributes from previous iterations impacts the fitness score of other remaining attributes, often resulting in a change of the order of the dimensions.

III. Datasets, Case Studies & Evaluation

We analyze three multidimensional datasets as part of this work. These datasets represent various services offered by a large network and service provider in the United States. Each dataset comes from different domains, related to network management and operation, but with different high-level behaviors.

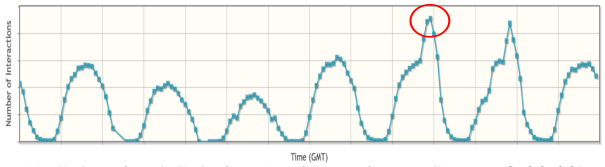
In this section we briefly describe each dataset, accompanied with case studies where we used AutoFocus to find the scope of real anomalies. These events were all confirmed via post-hoc analysis reports given by the operations team from the large network and service provider. We have anonymized individual attributes throughout the paper.

We also performed an evaluation using synthetic events. To inject a synthetic event into our real datasets we: (1) select a date and time, (2) randomly select a number of attributes to define the ground truth scope of the synthetic event, and (3) modify the observed values of the randomly selected attributes We used the median of the time series, multiplied by the length of the time series, as the observed value of the synthetic event. After injecting the anomaly, we ran AutoFocus and verified whether it found the target scope. To evaluate the success of the simulation, we checked to see if the results exactly matched the target. If there were any missed attributes, or if there were any extra, we count the simulation as a failure. We repeated the evaluation with 100 different randomly-generated synthetic anomalies.

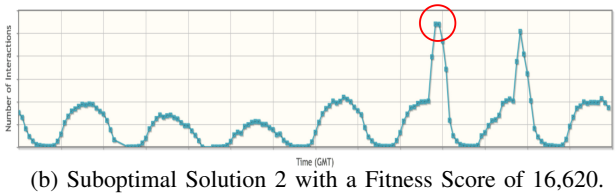
A. Customer Care Calls

This dataset consists of counters referring to the hourly number of customer calls placed to customer care centers reporting cellular service issues (e.g., dropped calls, slow video streaming). Typically the majority of customer calls concern individual customer issues, e.g., device issues or user errors. However, in some situations, these issues may be the result of a broader service impacting event. There are three dimensions in the customer care call data set: (i) Market: geographic region where the mobile user is located. (ii) Phone Model: model of the user's mobile phone. (iii) Application: the function operating on the phone such as SMS or VoLTE. An example entry in the dataset might be "12 mobile calls in Utah with Phone Model A using the VoLTE application". Our synthetic evaluation resulted in a success rate of 96.4% on this dataset.

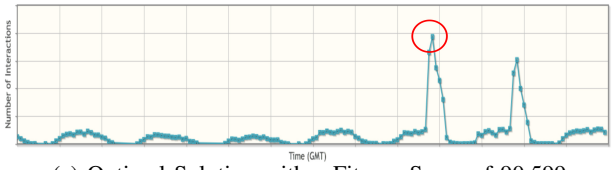
Case Study: Widespread Voice Call Drop Event. This case study finds an event with a noticeable increase in customer care calls concerning dropped voice calls. AutoFocus correctly identified the scope of the impact of this issue as confirmed by the network operations team Figure 2c shows a timeseries graph of the solution output by AutoFocus with the highest fitness score. The circle identifies the anomaly given as input to AutoFocus. Figures 2(a-b) illustrate two potential solutions evaluated by AutoFocus that were determined to be suboptimal. These time series graphs are taken from a tool used by the operations team to try and identify the scope of an event by hand. Using the fitness score, AutoFocus is able to automate the process for a quicker approach that is less tedious and error-prone.



(a) Suboptimal Solution 1 with a Fitness Score of 66,668.



(b) Suboptimal Solution 2 with a Fitness Score of 16,620.



(c) Optimal Solution with a Fitness Score of 90,599.

Fig. 2: Customer Care Case Study: Voice Call Drop Event. Each graph represents a time series of the customer care call dataset, filtered by different combinations of attributes. Values are removed for anonymization.

B. Cloud API Platform Error Logs

The next dataset consists of error logs collected from a large-scale operational cloud API platform providing various services (e.g., speech, payment, notary, SMS, advertisement) to third-party applications. Each API request first traverses a particular proxy server, and then interacts with a particular back-end server, as determined by load balancing. Error logs report any HTTP 4** and 5** error codes returned by these APIs. The number of these error codes are calculated from the logs and broken down into the following 3 dimensions: (i) API: provided by the cloud platform and used by any authorized users over the public Internet. (ii) Proxy Server: the façade of the API requests, handling authentication and load balancing. (iii) Back-end Server: where the business logic of various APIs are executed. An example entry in the dataset might be: there were 46 error log entries with HTTP status 504 on speech API requests going through proxy server 1 and back-end server 2.

Case Study: Issue with Multiple Proxies. This case study focuses on an event consisting of an increase in error responses from a few proxy servers. Since these proxies serve APIs used by third-party applications, it is essential to quickly identify exactly what proxies are having problems and which APIs are impacted by them. AutoFocus correctly identified the offending proxies as well as the APIs and back-end servers showing symptoms caused by the proxy issues.

C. LTE Traffic Performance Data

The last dataset is from an LTE mobile network. To ensure traffic is allocated appropriate Quality of Service (QoS) attributes such as delay and loss, a QoS Class Identifier (QCI) is used to prioritize traffic into QoS classes from the Radio Access Network (RAN) to the serving gateway inside the core network. When the traffic reaches the Packet Data Network (PDN) gateway, the Access Point Name (APN) is used to determine where to route the traffic. Each end-to-end traffic flow in the LTE network is associated with a QCI-APN pair, as well as various metrics (e.g., the number of dropped packets) measured for each flow. These three components of the LTE network are the dimensions we analyze in this paper: (i) QCI: The QoS Class Identifier for a set of traffic flows. (ii) APN: The APN being used by a set of traffic flows. (iii) PDN gateway: The PDN gateway that serves a geographic region.

We were not able to obtain any case studies for this dataset. However, the synthetic event evaluation resulted in a success rate of 88.6%.

IV. Related Work

The problem of impact scoping across multiple dimensions, to our knowledge, has not been widely studied. While anomaly detection has been studied extensively [4], impact scoping is the next step after anomaly detection. Specifically, consider previous works such as Argus [12], G-RCA [3], and ABSENCE [9], that focus on finding service-oriented events through event correlation on time series data. G-RCA tracks service dependencies across multiple datasets while Argus scopes them within within aggregate groups. ABSENCE is able to find otherwise undetected or "silent" events by correlating finding anomalies in customer usage volume.

RCATool [5] is an anomaly detection and diagnosis framework for Internet service anomalies, performing Root Cause Analysis on DNS traffic. The Distribution-based Anomaly Detection is closest to our work, using the entropy of an observed distribution of DNS requests compared to a reference distribution. When an anomaly is found, the elements experiencing impact from the anomaly can be determined by looking at their individual divergence over a certain threshold.

NetPoirot [1] uses machine learning to track patterns of TCP behavior in order to determine which network management team should be responsible for handling the trouble management process for a particular event. In a sense it localizes the scope of an event, but only enough to broadly describe the boundary within which the root cause of the event occurred. AutoFocus is more detailed in that it specifically identifies which individual components are associated with an event.

V. Conclusion

We presented AutoFocus, an algorithm to automate the process of finding the scope of the impact of anomalous events in multidimensional time series datasets. AutoFocus relies on its *fitness score* metric, which statistically defines the relevance of anomalous behavior. AutoFocus uses the fitness score to trim the search space of anomaly descriptions in order to find the scope of the impact of a given anomaly. We presented a data-driven, systematic evaluation of AutoFocus involving finding synthetic events in real datasets, as well as accurately detecting known ground truth events in real datasets, showing the feasibility of automated, accurate trouble isolation.

REFERENCES

[1] ARZANI, B., CIRACI, S., LOO, B. T., SCHUSTER, A., AND OUTHRED, G. Taking the blame game out of data centers operations with netpoirot. In *Proceedings of the 2016 Conference on ACM SIGCOMM 2016 Conference* (New York, NY, USA, 2016), SIGCOMM '16, ACM, pp. 440–453.

[2] BRAMLETTE, M. F. Initialization, mutation and selection methods in genetic algorithms for function optimization. In *ICGA* (1991), pp. 100–107.

[3] BRESLAU, H. Y. L., GE, Z., MASSEY, D., PEI, D., AND YATES, J. G-rca: A generic root cause analysis platform for service quality management in large ip networks.

[4] CHANDOLA, V., BANERJEE, A., AND KUMAR, V. Anomaly detection: A survey. *ACM Comput. Surv. 41*, 3 (July 2009), 15:1–15:58.

[5] FIADINO, P., DALCONZO, A., SCHIAVONE, M., AND CASAS, P. Rcatool-a framework for detecting and diagnosing anomalies in cellular networks. In *Teletraffic Congress (ITC 27), 2015 27th International* (2015), IEEE, pp. 194–202.

[6] GOLYANDINA, N., AND STEPANOV, D. Ssa-based approaches to analysis and forecast of multidimensional time series. In *proceedings of the 5th St. Petersburg workshop on simulation* (2005), vol. 293, p. 298.

[7] KENDALL, M. G., AND ORD, J. K. *Time-series*, vol. 296. Edward Arnold London, 1990.

[8] NABOULSI, D., FIORE, M., RIBOT, S., AND STANICA, R. Large-scale mobile traffic analysis: A survey. *IEEE Communications Surveys Tutorials 18*, 1 (Firstquarter 2016), 124–161.

[9] NGUYEN, B., GE, Z., VAN DER MERWE, J., YAN, H., AND YATES, J. Absence: Usage-based failure detection in mobile networks. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking* (New York, NY, USA, 2015), MobiCom '15, ACM, pp. 464–476.

[10] SHAFIQ, M. Z., ERMAN, J., JI, L., LIU, A. X., PANG, J., AND WANG, J. Understanding the impact of network dynamics on mobile video user engagement. In *ACM SIGMETRICS Performance Evaluation Review* (2014), vol. 42, ACM, pp. 367–379.

[11] WHITLEY, D. A genetic algorithm tutorial. *Statistics and computing 4*, 2 (1994), 65–85.

[12] YAN, H., FLAVEL, A., GE, Z., GERBER, A., MASSEY, D., PAPADOPOULOS, C., SHAH, H., AND YATES, J. Argus: End-to-end service anomaly detection and localization from an isp's point of view. In *INFOCOM, 2012 Proceedings IEEE* (March 2012), pp. 2756–2760.