

Fehler-Erkennung in Netzwerken

Die Betreiber von Netzwerkdiensten müssen Anomalien in ihren Systemen möglichst frühzeitig erkennen, um entsprechende Gegenmaßnahmen einleiten zu können. Klassische Methoden, etwa die Nutzung fester Grenzwerte, sind dazu zwar geeignet – sie sind jedoch nur statisch. Anders die Methode zur automatischen Anomalieerkennung, die auf der Auswertung mehrdimensionaler signalbasierter Messwerte beruht.

Die Methode passt sich automatisch an allmähliche Änderungen und periodische Last-Zyklen der gemessenen Signale an und unterstützt die Auswertung mehrerer unabhängiger Messwerte. Das Ergebnis: Analysen mit einer deutlich verbesserten Aussagekraft. Denn die neue Methode trifft nicht nur qualitative (ja/nein) Aussagen darüber, ob eine Anomalie vorliegt, sondern quantifiziert diese Aussage zusätzlich mit einem Zahlenwert (dem so genannten Konfidenzwert), der angibt, wie sehr die gemessenen Signale auf eine Anomalie hindeuten.

Leistungsüberwachung und Anomalieerkennung sind für komplexe Netzwerkdienste von immenser Bedeutung. Sie hel-

stark und unvorhergesehen wechselnde Lastsituationen bewältigen. Eine statische Festlegung von Grenzwerten muss also zwangsläufig scheitern.

Eine Analyselösung muss aber Angriffe und Fehler in der eigenen Infrastruktur (zum Beispiel Hardwareausfälle oder Konfigurationsprobleme) von harmlosen oder gewünschten Lastspitzen unterscheiden können. Flash-Crowd-Ereignisse – etwa ein durch die Medien in kürzester Zeit populär gewordener Netzwerkdienst – bewirken, dass die Last auf den Systemen rapide ansteigt, was in diesem Fall natürlich gewünscht ist. Falls dieser plötzliche Erfolg jedoch fälschlicherweise als Denial-of-Service-Angriff (DoS) missinterpretiert wird, können die dann ergriffenen Schutzmaßnahmen Probleme verursachen. Im schlimmsten Fall kann die Reaktion auf den vermeintlichen Fehler zur Nichterreichbarkeit des Dienstes führen, obwohl die Serverkapazität den Nutzer-Ansturm eigentlich hätte bewältigen können.

Auswertung mehrdimensionaler Messwerte

Der neue methodische Ansatz ermöglicht die Erkennung von Netzwerkproblemen und vieler Denial-Of-Service-Angriffe mit hoher Zuverlässigkeit und insbesondere deren Unterscheidung von Flash-Crowd-Ereignissen. Dazu werden signalbasierte Messwerte (etwa eingehende

und ausgehende Datenmengen, CPU-Last, Speicherverbrauch, Anzahl der Anfragen pro Sekunde) und ihre Veränderung über die Zeit betrachtet.

Kann das System die Anfragen problemlos bearbeiten, steigt der Durchsatz unge-

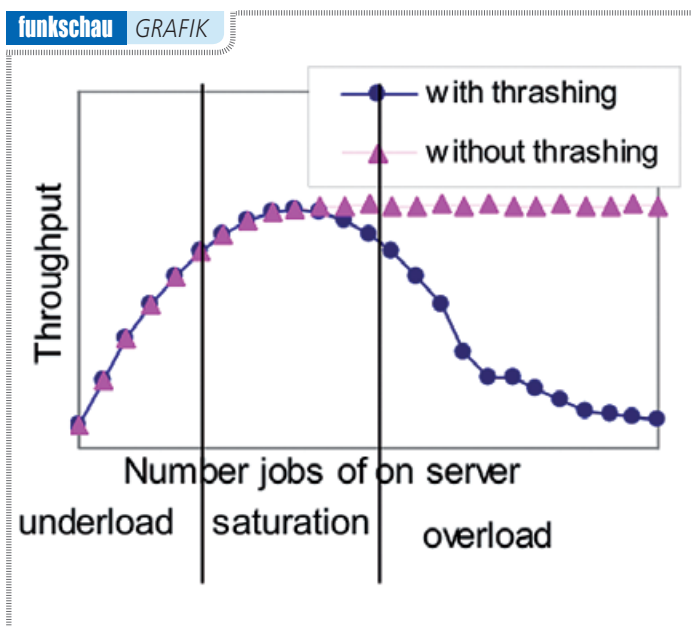
fähr linear mit der Anzahl der Anfragen. Stößt das System auf Grund weiter steigender Last an seine Leistungsgrenzen, können jedoch zwei mögliche Szenarien auftreten:

Im ersten Szenario bleibt der Durchsatz in etwa konstant (die so genannte Sättigungsphase): Der Server bedient die einzelnen Anfragen mit Verzögerung, arbeitet aber weiterhin gut. Je nach Ursache des Leistungslimits oder bei weiter steigender Last kann jedoch der so genannte Trashing-Effekt auftreten. Dann bricht der Durchsatz des Systems rapide ein. Dieser Fall tritt zum Beispiel auf, wenn dem Server zu wenig RAM zur Verfügung steht und er ab einer bestimmten Belastung auf langsamen Swap-Speicher ausweichen muss.

Auch bei bestimmten Arten von DoS-Angriffen steigt die Last auf dem Server sehr stark an, während der Ausgangstraffic allerdings relativ niedrig bleibt. Auf diese Weise sind selbst Angreifer mit durchschnittlichem Netzwerkzugang in der Lage, bestens angebundene Server von großen Anbietern anzugreifen. In beiden Szenarien ist es hilfreich, nicht nur einen einzelnen Messwert (zum Beispiel die CPU-Last des Servers) für die Analyse zu überwachen, sondern mehrere gleichzeitig zu beobachten und in Zusammenhang zu bringen (eine so genannte mehrdimensionale Analyse).

Referenz- und Beobachtungsfenster

Um adaptive und quantifizierbare Resultate zu erhalten, wird ein Ansatz aus dem Bereich der Lern-Algorithmen angepasst und der oben erwähnte Konfidenzwert berechnet. Dazu werden die in einem Referenz- und einem Beobachtungsbereich gemessenen Signale verglichen. Die aktuellen Daten im Beobachtungsbereich werden dabei mit einem Profil normaler Messwerte aus dem in der Vergangenheit liegenden Referenzbereich verglichen. Beide Bereiche bewegen sich dabei mit der Zeit synchron weiter.



Das Bild zeigt schematisch einen typischen Lastverlauf bei einem stark beanspruchten Server.

fen Angriffe und Probleme rechtzeitig zu erkennen und zu beheben. Die Anpassung von herkömmlichen statischen Überwachungsmethoden ist jedoch nur begrenzt für komplexe, verteilte Dienste geeignet. Der Grund: Diese Dienste müssen häufig

Die Länge beider Bereiche muss dabei nicht identisch sein. In der Regel umfasst der Referenzbereich einen deutlich größeren Zeitraum (zum Beispiel einen Tag) als der Beobachtungsbereich, der meistens nur wenige Messwerte umfasst. Damit Anomalien gut erkannt werden können, sollte das Referenzfenster zum Zeitpunkt einer Anomalie ausschließlich „normale“ Werte enthalten.

Beide Bereiche werden durch die Berechnung der so genannten Mahalanobis-Distanz miteinander in Beziehung gebracht. Sie gibt wieder, wie stark sich die Signale in ihnen unterscheiden. Die Mahalanobis-Distanz ist eine effizient berechenbare mathematische Klassifikationsmethode und kann zum Vergleich mehrdimensionaler Daten angewandt werden. Dabei muss nicht explizit angegeben werden, wie die verschiedenen Dimensionen zusammenhängen.

Die Mahalanobis-Distanz verhält sich dabei nicht linear zur geometrischen Distanz (im n -dimensionalen Raum) beider Gruppen, sondern berücksichtigt die Verteilung der Messpunkte in den jeweiligen Gruppen. Außerdem ist sie skalunabhängig, das heißt nur die relativen Abweichungen werden für die Distanz berechnet – der absolute Wert oder die Einheit von Messwerten sind irrelevant.

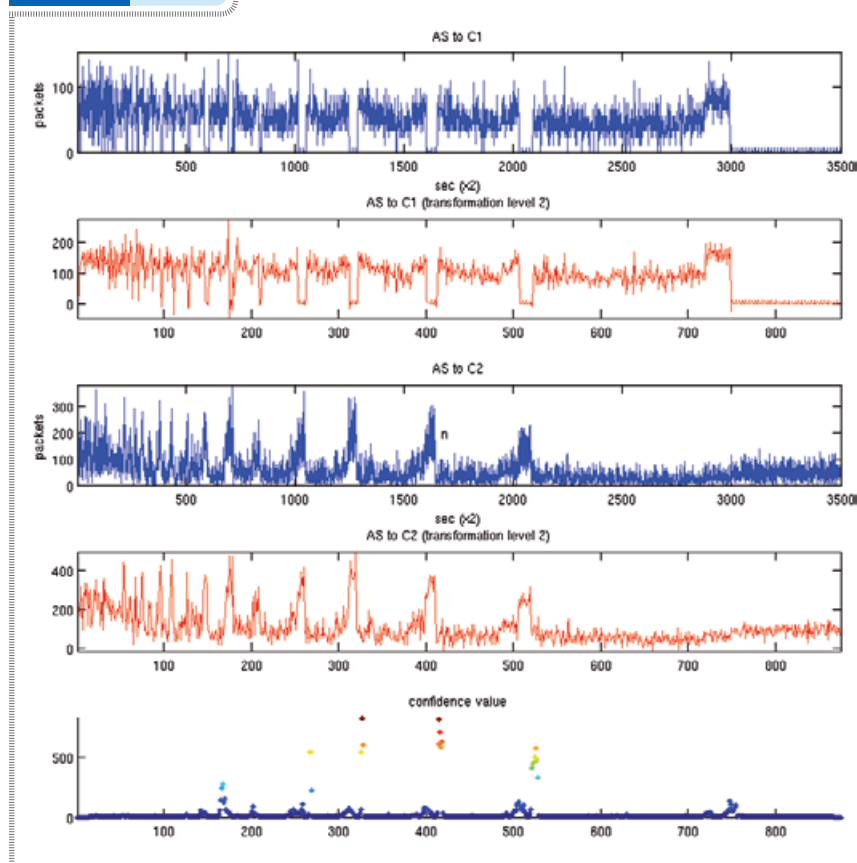
Rauschminderung ist notwendig

Eine wesentliche Voraussetzung für die Nutzung der Mahalanobis-Distanz sind sehr rauscharme Signale. Bei den verwendeten Messwerten ist dies aber in aller Regel nicht der Fall. Die Werte müssen also zunächst geglättet werden. Dabei kommen Wavelet-Transformationen zum Einsatz, die sehr effizient durchgeführt werden können und für unseren Anwendungsfall gute Ergebnisse mit einem hohen Signal-Rausch-Verhältnis liefern.

Das erreichbare Glättungsniveau ist durch die Sampling-Rate der gemessenen Signale beschränkt, denn jede Iteration der Wavelet-Transformation halbiert die Auflösung eines Signals. Dass die Wavelet-Transformation die Messwerte oft nach oben skaliert, stellt für die hier vorgestellte Methode jedoch kein Problem dar, da die Mahalanobis-Distanz wie erwähnt skalunabhängig arbeitet.

Die vorgestellte Methode kann durch Veränderung verschiedener Parameter je nach Szenario und verfügbaren Ressourcen angepasst werden. Dabei spielen insbesondere die Stufe der Wavelet-Transformation und die Länge von Beobachtungs- und Referenzfenster sowie deren zeitlicher Abstand eine große Rolle. Ein kürzerer Abstand zwischen den Fenstern bedeutet, dass das System Änderungen schneller lernt (da sie früher im

funkschau GRAFIK



Bilder: Consistec

Die Abbildung zeigt die Anwendung der vorgestellten Methode an Messwerten eines realen Netzwerkdienstes, bestehend aus einem Application-Server (AS) und zwei Backend-Servern C1 und C2.

Referenzfenster ankommen) und somit weniger anfällig für Fehlalarme bei schleichenden Änderungen der Signale ist. Ist der Abstand jedoch zu kurz, wird das Referenzfenster unter Umständen mit anomalen Werten kontaminiert. Anomalien bleiben dadurch unter Umständen unbemerkt.

Durch die Wahl der Länge des Referenzfensters kann Einfluss darauf genommen werden, wie gut das System regelmäßige Schwankungen der Messwerte toleriert. Enthalten Signale etwa bestimmte Nutzungsmuster, zum Beispiel tageszeitabhängige Schwankungen des Netzwerkverkehrs, sollte das Referenzfenster mindestens eine komplette Periode (zum Beispiel einen Tag) umfassen. Das System kann so die natürlichen Schwankungen optimal berücksichtigen.

Auswahl der analysierten Messwerte

Signale zur Beobachtung auszuwählen, die sich im Falle einer Anomalie alle ähnlich (redundant) verhalten, ist wenig sinnvoll. Denn eine wichtige Eigenschaft der Methode ist es, Anomalien in mehreren (unabhängigen) Dimensionen zu erkennen.

Mehrdimensional bedeutet hier, Anomalien nicht nur in einzelnen Signalen zu orten, sondern auch das abnormale Verhalten eines Signals in Bezug auf die anderen Signale. In der Praxis sind genau diese Phä-

nomene für die Zustandsüberwachung eines Systems hochinteressant: So kann beispielsweise die Eingangsrate von Nutzeranfragen und die Antwortrate eines Netzwerkdienstes analysiert werden, um die Lastkapazität zu überwachen oder auch um Flash-Crowd-Ereignisse von DoS-Angriffen zu unterscheiden. Steigt die Requestrate rasch während die Antwortrate deutlich fällt (wegen DoS-Angriffen oder Überlast), wird dabei ein deutlich größerer Konfidenzwert berechnet als bei einem Anwachsen der Antwortrate entsprechend der Requestrate (Flash Crowd).

Ein Beispiel: Die obige Grafik zeigt die Anwendung der vorgestellten Methode an Messwerten eines realen Netzwerkdienstes, bestehend aus einem Application-Server (AS) und zwei Backend-Servern C1 und C2. Dabei wurde als Messwert jeweils die Anzahl der Pakete pro Sekunde zwischen AS und den Backend-Servern gewählt. Diese wird einmal vor der Wavelet-Transformation (ungeglättet) und einmal danach dargestellt. Im untersten Diagramm ist schließlich der berechnete Konfidenzwert dargestellt. Sehr hohe Konfidenzwerte weisen deutlich auf potenzielle Anomalien hin. (WM)

 **Kurnia Hendrawan**
Software Engineer bei Consistec