

Pay Attention to Anomalies

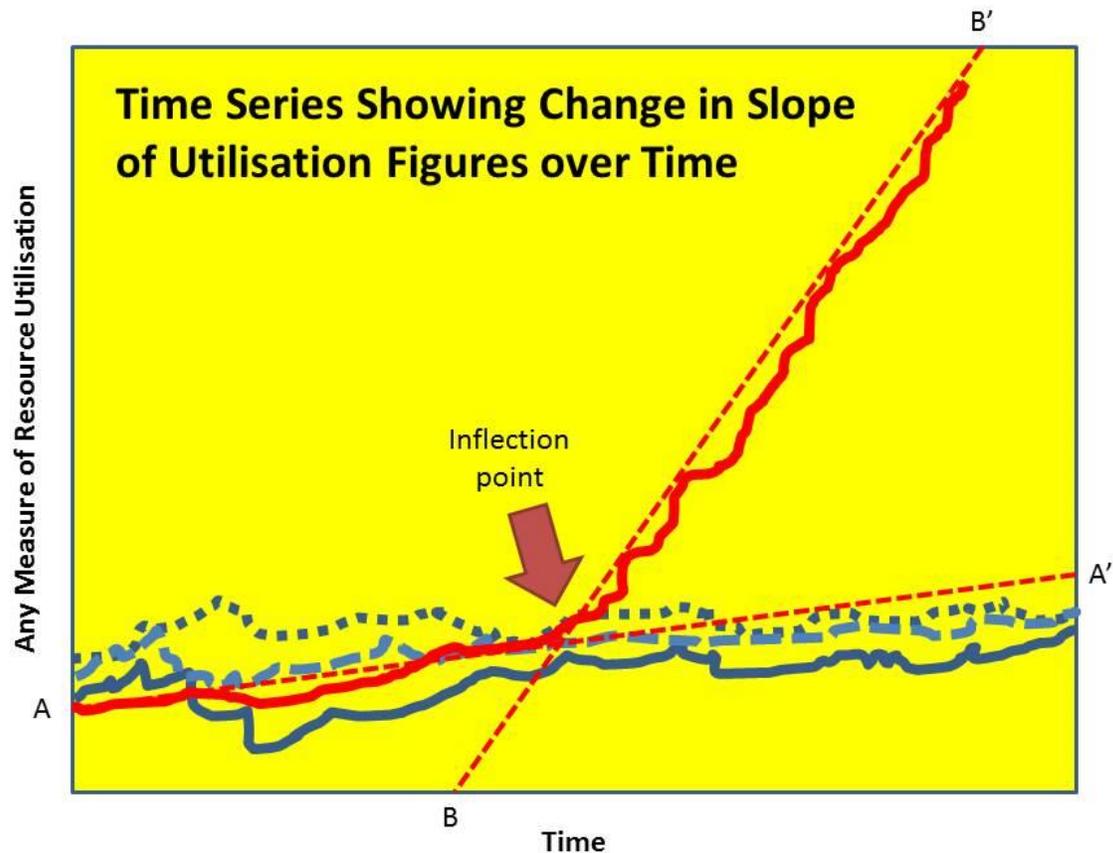
by M. E. Kabay, PhD, CISSP-ISSMP
Professor of Computer Information Systems
School of Business & Management
Norwich University, Northfield VT

Today I increased my virtue coefficient by getting to the swimming pool up the road from where I live (well, 7 km from where I live in farming country) early in the morning. On my way out after a vigorous set of laps (I normally swim a “mile,” which is an ancient measure of distance still used in backwaters such as the USA), I stopped at the desk to tell the attendant that I would like to switch my automatic payments from my credit card to a direct withdrawal from my bank account (VISA charges are rough on the profits of this small business in the wilds of Vermont and I’d like to do my part to help these folks out).

Jim the attendant looked me up on his computer and discovered that there had been no payments from my VISA account since last November: I owed six months of fees! Jim explained that there had been “some problems” with their admission and billing system after an upgrade in January, but that the problems were resolved now. He agreed that allowing members to default on their monthly dues was a serious threat to cash flow for this small organization. It seems that the software was *supposed* to issue a warning automatically to members about delinquent payments, but that part of the code never activated.

After paying the accumulated fees, I thought about how dangerous this glitch – this anomaly – was. This health club has a sign-in that uses a card swipe to bring up the members account on screen and shows a photo of the authorized user of the account – all in aid of user identification and authentication to prevent free riders who might share a single membership. In previous discussions with managers, I’ve urged them to monitor their records carefully for customer-relationship assurance; for example, if I were running the club, I’d bring up an exception report whenever any member’s attendance figures dropped below a proportion of their normal activity. Someone who works out every day hasn’t shown up for a week: are they ill? Are they on a trip or on vacation? An e-mail expressing concern would be a nice touch; perhaps the managers would respond to a sickness by offering the customer an extension on their account – you don’t use the system for a month, you get the deadline extended by a month. Note that the critical issue in such a system is to define the anomaly in terms of what’s normal for the specific member, not in terms of a general average. Thus if someone comes in once a week and misses a month, then perhaps *that* situation would prompt an enquiry.

I’ve taught students for decades that they should be paying attention to anomalies. Anomaly detection depends on adequate data gathering and statistical analysis – or even just graphical representation of the data. For example, the following diagram shows a typical anomaly: a change of slope.



In the diagram, the three blue lines at the bottom represent some normal changes in a resource; e.g., disk-space utilisation. The graphic reminds me of a good real-world illustration of the value of paying attention to anomalies.

Back in the mid-1980s (my goodness, almost 30 years ago!), I was the director of technical services for a computer time-sharing service bureau (we served 28 insurance companies and insurance brokerages) in Montréal, Québec, Canada. I remember a specific incident where I was using a graph with lines for disk space utilisation (when a huge HP7933 <<http://www.hpmuseum.net/images/7935-40.jpg>> disk drive had 404MB – yes, megabytes, children, not terabytes) by different customers. In the case I am remembering, a customer's disk-space began growing at a furious rate (represented as the continuous red line in the diagram). The *inflection point* is where the slope in disk utilisation changed from the A-A' line to the B-B' line. Any time there's an inflection point, resource managers should become curious: what's happened to cause the inflection? Because I routinely monitored CPU utilisation and disk-space utilisation among other parameters, I spotted the change quickly and investigated. It turned out that programmer at the customer site had REMmed (commented out – from REMark) the instructions in the job control language (JCL) for a particular batch job so that temporary files wouldn't be deleted at the end of the job. After he ran his diagnostics, he forgot to remove the REMs. By the time I caught the anomaly, the client had about 20,000 temp files in their account. (By the way, that programmer should *not* have been using *production* code to run his tests.)

Even if a change in slope is not due to an error, noticing and investigating an inflection point is a good idea. For example, it could be that a new routine has been implemented on the date where

the slope has changed; in that case, system managers would want to notice the change in resource consumption and plan for orderly resource management (such as ordering new resources earlier or later than planned).

From an *information security* perspective, inflection points in resource utilisation can signal information system security officers (ISSOs) that something unusual has happened or that a norm is being redefined. For example, suppose that the graph above represented accumulated CPU utilisation or bandwidth utilisation for individuals or for specific workgroups in an organization. Wouldn't any ISSO want to know *why* there was a change? What if it were an unauthorized change? What if the system had been infected by botnet malware and the increased bandwidth was due to 10,000 spam e-mails being sent out per hour on a rogue Simple Mail Transport Protocol (SMTP) server<
http://www.blockdos.net/mind_the_egress_when_filtering_for_security.html>?

Another example: Joe the accountant has never logged into the network after working hours in the last six years of record-keeping; so what is happening when "he" starts logging in at 03:00 every day and is generating GB of data transfers? Wouldn't an ISSO want to check with Joe about what's up? And wouldn't it be important to discover that Joe has no idea what the ISSO is talking about? Aha! Unauthorized access: hacker at work.

Anomaly detection using resource utilisation data can't be invoked suddenly: unless there are accumulated data allowing analysts to establish norms, it may be difficult or impossible to distinguish random fluctuations from systematic changes. For those interested in automating their analytical tools, one would compute linear or nonlinear regression coefficients for moving subsets defined by some reasonable period (as a function of the intrinsic variability of the data) and note changes automatically for alerts to be signalled to the resource managers or ISSOs. Readers will want to consult any textbook of applied statistics for details.

So to sum up, keep track of resource utilisation and investigate anomalies!

* * *

M. E. Kabay,<mailto:mekabay@gmail.com> PhD, CISSP-ISSMP, specializes in security and operations management consulting services and teaching. He Professor of Computer Information Systems in the School of Business and Management at Norwich University. Visit his Website for white papers and course materials.<http://www.mekabay.com/>

* * *

Copyright © 2012 M. E. Kabay. All rights reserved.