Short-term Periodicity in Security Vulnerability Activity

HyunChul Joh, Sopak Chaichana and Yashwant K. Malaiya
Computer Science Department
Colorado State University
Fort Collins, CO, USA
{dean2026, chaichan, malaiya}@cs.colostate.edu

***Abstract*—Some of the major computer security organizations monitor a global pool of systems for presence of vulnerabilities and worms. The extensive amount of data generated provides important insights into the vulnerability activity and the risk they represent. An examination of the data published suggests weekly periodical behavior. This paper identifies the periodicity and examines its statistical significance for some of the data series published. The results shows that the seven-day periodicity in presence of unpatched vulnerabilities as well as the exploitation pattern. The behavior during the weekdays itself is found to vary. This behavior should be used to optimize resource allocations and for determination of risk.**

Keywords-Software vulnerability; Autocorrelation fucntion analysis; periodic behavior; seasonal index; weekend effect

I. INTRODUCTION

Periodic scanning is a major part of the corporate security strategy. Some security service venders like Qualys collect a large amount of data which is quite valuable because it comes from real systems in major industrial organizations. In this study, we have mined one such data collection to examine periodicity in the presence of unpatched vulnerabilities and the exploitations in case of a worm.

Qualys has been involved in collecting and plotting such data for several years. In a 2009 report, they have presented the data collected during 2008 which represents 104 million global vulnerability scans including 82 million internal scans and 22 million external Internet-based scans. The data involves encountering more than 72 million critical vulnerabilities among the 680 million detections. About 3500 organizations were scanned worldwide that represented major industry sectors of Financial, Health, Manufacturing, Service, and Wholesale/Retail.

We observed that most of the plots, in the Qualys report [1], visually suggest a short-term periodicity. This paper presents two selected data series, shown in Fig. 1, from the report to mathematically examine the periodicity using statistical methods. The autocorrelation function analysis (ACF) is utilized to identify the exact periodicity. Also, the statistical significance of the periodical pattern for the two plots is examined by using seasonal index analysis with the χ^2 test. The same methods were used by Joh and Malaiya [2] for the longer term periodicity in the vulnerability reporting patterns for a dozen major software systems. While all the programs exhibit a year-end peak, additional higher incidences are also observed during the mid-year months for Microsoft products.

Some related work justifies the need for this research. Anbalagan and Vouk [3] suggest a possible weekly pattern for fixing of ordinary defects in Ubuntu 8.04.

II. ANALYZING WEEKLY PERIODICAL BEHAVIOR

We used the ACF method to determine the periodicity. ACF analysis helps to specify a relationship between the related time intervals. If the time series is given by z_b, z_{b+1}, …, z_n, the ACF at time lag k, denoted by r_k , is [4]:

$$r_k = \frac{\sum_{t=b}^{n-k}(z_t-\bar{z})(z_{t+k}-\bar{z})}{\sum_{t=b}^{n}(z_t-\bar{z})^2} \text{ , where } \bar{z} = \frac{\sum_{t=b}^{n} z_t}{(n-b+1)} \quad (1)$$

ACF analysis measures the linear relationship between time series observations separated by a lag of k time units. Hence, if measured ACF values are located outside of chosen upper or lower confidence intervals, the periodicity associated with those time lags is statistically significant.

We will examine the null hypothesis that no periodicity is present using the seasonal index analysis [5] which measures how much the average for a particular period tends to be above (or below) the expected value. The daily seasonal index values are given by:

$$s_i = \frac{d_i}{d} \quad (2)$$

where, s_i is the seasonal index for i^{th} day of the week, d_i is the mean value of i^{th} day (Monday ~ Sunday), d is a grand average. For example, a seasonal index of 1.2 indicates that the expected value for that day is 20% greater than 1/7 of the overall average where the expected value is 1. The χ^2 test for the null hypothesis has been conducted to check the statistical significance.

III. RESULTS

The upper two plots in Fig. 1 show the presence of the unpatched critical vulnerabilities and the detected number of exploitations of MS08-067 (Windows OS related vulnerability) by the Conflicker worm with χ^2 statistic values, χ^2 critical values and p-values. Incidentally, the Qualys report uses the first time series to determine the *half-life* i.e. the time it takes for organizations to remedy half of their vulnerable systems. The plots are normalized using the maximum value as 100%. Even visually, it is clearly observed that there are certain periodical patterns in the data. The values decline as the result of a remediation and go up due to new installations.

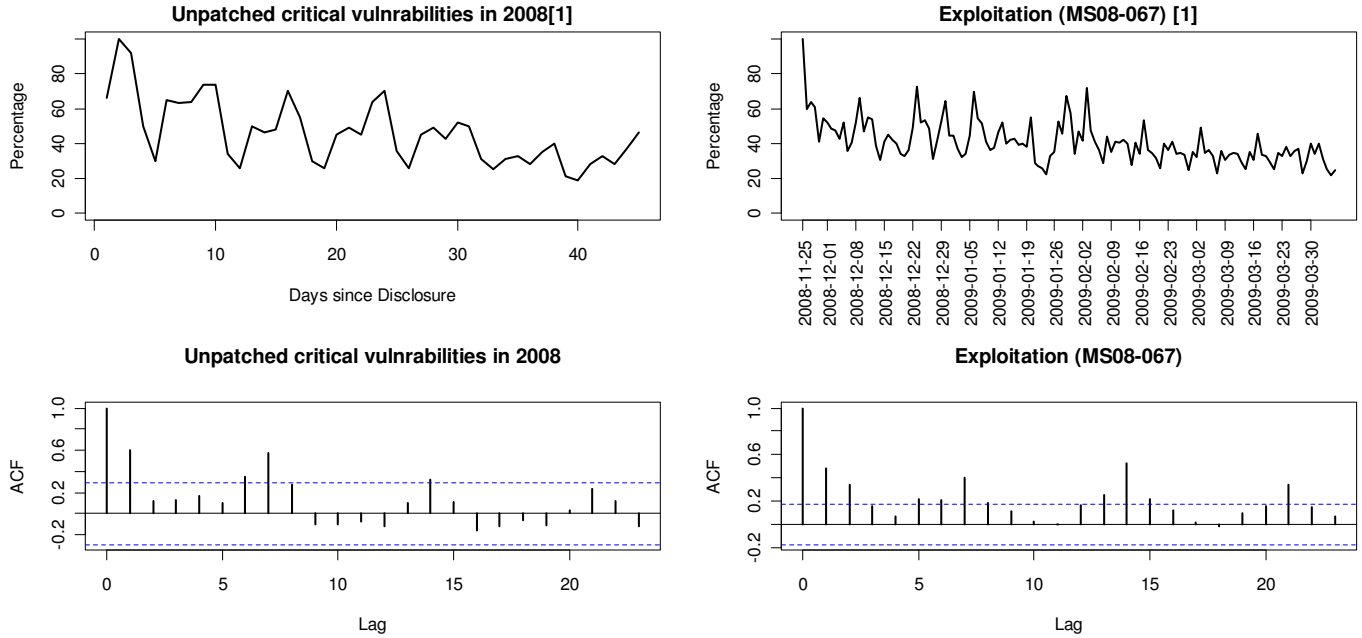

Figure 1. Run charts for unpatched critical vulnerabilities in 2008 and Exploitation with their corresponding ACFs. The upper two plots are normalized using the maximum value as 100%. In the bottom two plots, legs are in day.

TABLE I. SEASONAL INDEX VALUES

		Unpatched Vuln. (critical in 2008)	Exploitation (MS08-067)
Mon	day1	1.0495	1.0069
Tue	day2	1.4100	1.2973
Wed	day3	1.3600	1.0203
Thu	day4	0.7211	1.0354
Fri	day5	0.5426	0.9534
Sat	day6	0.9424	0.7307
Sun	day7	0.9745	0.9560
χ^2 critical		12.5916	12.5916
χ^2 statistic		165.6114	119.9789
p-value		3.83E-33	1.65E-23

The bottom two plots from Fig. 1 show the corresponding ACFs for the data series. Both of them have lags of seven (or its multiples) outside of the $\pm$ 95% confidence intervals shown by the dotted lines. This demonstrates strong autocorrelations with lags that are multiples of seven days, which confirms a seven-day periodicity in the data.

TABLE I shows the calculated daily seasonal index values for the two data sets. Since there is no information for day of the week in the first plot, it is tabulated as day1, day2, …, day7 while the exploitation data mentions specific weekdays (Monday through Sunday).

It is clearly seen that weekdays (Mon~Thu) tend to have higher index values for the number of incidents. The seasonal index value for the number of incidents for MS08-067 hits the highest on Tuesday. For the column for unpatched vulnerabilities, day1~day3 tend to have higher values than day4~day7. This might be related to the software vendors' policies [6] or individual behavior [7]. To be statistically significant for the calculated seasonal index values, the χ^2 statistic values need to be greater than the χ^2 critical value with a small enough p-value. In the table, the small p-values confirm the non-uniform distributions.

The results show the strong weekly periodicity in presence of unpatched vulnerabilities and the number of exploitations for MS08-067 with higher activities during the specific days of the week. Thus this periodicity should be kept in mind to optimize resource allocation.

REFERENCES

[1] W. Kandek, The Laws of Vulnerabilities 2.0, *Black Hat 2009*, July 28, 2009. Available at web:
-Report: http://www.qualys.com/docs/Laws_2.0.pdf
-Slides: http://www.qualys.com/docs/Laws_2.0_PPT.pdf

[2] H. Joh and Y.K. Malaiya, Seasonal Variation in the Vulnerability Discovery Process, *2nd International Conference on Software Testing Verification and Validation*, pp.191-200, 2009.

[3] P. Anbalagan, and M. Vouk, "Days of the week" effect in predicting the time taken to fix defects. *2nd international Workshop on Defects in Large Software Systems,* pp. 29-30, 2009.

[4] B. L. Bowerman and R. T. O'connell, *Time Series Forecsting*: *Unified concepts and computer implementation*.2nd Ed., Boston: Duxbury Press, 1987.

[5] H. Arsham. Time-Critical Decision Making for Business Administration. Available at web: http://home.ubalt.edu/ntsbarsh/Business-stat/stat-data/Forecast.htm#rseasonindx

[6] Microsoft Security Bulletin Advance Notification. Available at web: http://www.microsoft.com/technet/security/bulletin/advance.mspx

[7] Tufin Technologies, Tufin Survey: Hackers Say Take a Break This Summer Before Winter Hacking Spike, 2009 Available at web: http://www.tufin.com/news_events_press_releases.php?index=2009-08-25