Temporal Effects of Contributing Factors in Insider Risk Assessment: Insider Threat Indicator Decay Characteristics

Frank L. Greitzer†
PsyberAnalytix
Richland WA USA
Frank@PsyberAnalytix.com

Rex A. Kliner
Department of the Air Force
San Antonio, TX USA
rex.kliner.1@us.af.mil

Samantha Chan
Cogility Software
Irvine, CA USA
schan@cogility.com

ABSTRACT

This paper presents initial research results for a study examining temporal effects on the impact of insider threat indicators in judgments of insider threat risk. Earlier work focusing on a small set of potential risk indicators (PRIs) suggests that insider threat indicators may have varying temporal effects on judgments of insider risk. The present study sought to obtain judgments of temporal effects for a large set of insider threat risk indicators. Results of an expert knowledge elicitation study conducted to examine possible temporal effects suggested that PRIs vary in their temporal effects on expert judgment of insider risk and that there are systematic differences in PRI "half-life" based on indicator characteristics. Implications of these findings are discussed.

CCS CONCEPTS

• Security and privacy • Human and societal aspects of security and privacy

KEYWORDS

insider threat, insider risk, threat assessment, risk assessment, cybersecurity, insider threat indicators, information security

ACM Reference format:

Frank L. Greitzer, Rex A. Kliner, and Samantha Chan. 2022. Temporal Effects of Contributing Factors in Insider Risk Assessment: Insider Threat Indicator Decay Characteristics. In *Proceedings of the Annual Computer Security Applications Conference (ACSAC), Workshop on Research for Insider Threat (WRIT), December 5, 2022, Austin, TX USA. 10 pages.*

1 Introduction

Insider threats are people with access to an organization's information or assets (facilities, resources, people) who act in ways that may harm the organization [1]. This paper examines temporal effects that may influence a threat assessment expert's judgment of insider risk. The contribution of this paper is to further inform the argument that insider risk assessment should transcend the simple "accounting" of the intrinsic threat/risk level of observed individual indicators [2].

Research and common sense suggest that in addition to the "severity" of any insider threat indicator, the passage of time since observing the indicator should be considered when assessing insider threat risk. However, the effects of time elapsed since the observation or reporting of a Potential Risk Indicator (PRI) has not been systematically examined. Specifically, we ask if the impact of a PRI should be expected to remain constant over time, or should we consider the possibility that its effect on insider risk judgment may diminish (decay)? Further, may we assume that all PRIs undergo similar rates of decay? For example, a current case exhibiting multiple failed authentication attempts might be evaluated as far more concerning than the same behavior occurring in the distant past, with no intervening compliance violations. On the other hand, an analyst may continue to consider some indicators even though they occurred in the distant past (such as hostile acts, bullying/harassment, or criminal activity).

Unlike many studies of decay processes that focus on physical systems that decay or fail over time, here we address the hypothetical "influence" of an observed PRI on the analyst's judgment of insider risk, as a function of time. We refer to this temporal factor as indicator decay and we wish to estimate decay rates. If the influence of a PRI on the threat analysis process decreases over time, we can describe this temporal process in mathematical terms, such as by using a linear or an exponential decay function requiring only a single decay parameter; and we can characterize the process further by indicating its half-life—the length of time that it takes for the impact of a PRI to decrease by fifty percent.

This paper is organized as follows: Section 2 discusses related research that helps to inform our expectations about

†Corresponding author

ACSAC WRIT Workshop, December, 2022, Austin, Texas USA

the nature of PRI decay. Section 3 describes the possible effects of decay in mathematical terms. Section 4 presents the research hypotheses. Section 5 describes the method and procedure of an expert knowledge elicitation study, prior to more formal studies that are planned, to identify insights that might be gained about temporal effects. In Section 6 we describe the results. Section 7 discusses what we learned and implications for the conduct of planned studies. Section 8 concludes the paper with a brief discussion of limitations and implications of this research.

2 Research on Insider Threat Indicators and Decay Characteristics

2.1 Insider Threat Indicators

Several researchers have developed models or frameworks describing individual insider threat indicators (e.g., [3] [4] [5]), but few ([6] [7] [8] [9]) have attempted to quantify or distinguish among the contributions of different indicators to possible insider threat risk. Two of the most robust sources of insider threat indicators are the ontologies developed in [7] and [9]. The Sociotechnical and Organizational Factors for Insider Threat (SOFIT) ontology [7], which comprises nearly 300 individual and organizational indicators, specifies five major classes of PRI: *Boundary Violation, Cybersecurity Violation, Job Performance, Life Narrative Factors,* and *Psychological Factors* (the full ontology is available to download in [10]). Another source of insider threat indicators is a list of PRIs developed by the Department of Defense Insider Threat Management and Analysis Center (DITMAC), comprising about 136 PRIs organized into nine categories: *Access Attributes; Professional Lifecycle and Performance; Foreign Considerations; Security and Compliance Incidents; Technical Activity; Criminal, Violent, or Abusive Conduct; Financial Considerations; Substance Abuse and Addictive Behaviors;* and *Judgment, Character and Psychological Considerations.* Both the SOFIT and DoD taxonomies address technical and behavioral/psychological factors and, therefore, these represent the most comprehensive knowledge bases that are currently available. We base our current study on these two information sources.

2.2 Relevant Research

There is not a great deal of relevant research to draw upon for insight on this research question. We may surmise that the time span over which an analyst should take PRIs into account can be informed by research focusing on consumers, or in other contexts, staff members, who feel aggrieved after having negative experiences. Studies of customer grudges conclude that grudge holders are emotionally upset at the incident that produced the grudge (which we may think of as a Precipitating Event in our research context); they remain upset over time, often for years or even decades [11]. For customers who feel they have been wronged in some way, desire for revenge gradually decreases and desire to avoid (not shop with) the company increases over time since a Precipitating Event occurred that caused this sense of grievance [12]. For loyal customers, Gregoire and colleagues [13] [14] found that the desire for revenge decreases more slowly over time and the avoidance response increased more rapidly. Feelings of betrayal—which was described as a key predictor of a desire for revenge—decreased nonlinearly over time. Inspection of their results [13] strongly suggests that the decay function is not linear—it more resembles an exponential decay. According to their research, the desire for revenge decreased by about 25% over the eight week duration of their study (~60 days, with measurements obtained in 2-week intervals over eight weeks).

In the only study we are aware of addressing possible decay of insider threat indicators, experts were asked to judge "level of concern" or severity for PRIs over time [2]. Only about ten percent of the indicators listed in the SOFIT ontology were studied, with PRIs representing four Role Types: *Personal Predisposition, Precipitating Event, Behavioral Precursor,* and *Technical Precursor*. The results suggested that all the PRIs exhibited some decay; the decay functions roughly resembled a linear decrease (at least for the limited timespan studied of 3-4 months), with a monthly decrement of about 8% of the PRI's initial severity value, which corresponds to a decrement of 0.08/30 = 0.00267 per day. However, a closer analysis of the data reported revealed that *Personal Predispositions* were about one-half as likely to show any decrease at all, from month to month, compared to the other Role Types. Of the 46 cases involving *Personal Predispositions* (*Manipulative, Big Ego*), 18 cases (39%) were judged to show some decay in severity, but 28 of these case judgments (61%) showed no decay across the four months that were rated—thus, significantly, nearly two-thirds of the threat ratings for the *Personal Predispositions* were *stable* over time. In contrast, of 45 cases involving Technical Precursors (*File Deletion, Unauthorized Database Searches*), 36 (80%) were judged to decay in severity while only nine (20%) showed no decay across the four months. For the 45 cases involving *Behavioral Precursors* (*Disgruntled, Anger/Hostility*), eight cases showed no decay and 37 (82%) were judged to decay over time. For *Precipitating Events* (*Corrective Action, Terminated*), 40 of 45 cases (89%) represented decreases in threat ratings. Thus, for

Technical Precursors, *Behavioral Precursors*, and *Precipitating Events*, 80-90% of the threat ratings showed some level of decay over time, in sharp contrast to *Personal Predispositions*. These variations in the likelihood of indicator threat ratings to decay were reported to be highly statistically significant [*Chi-square (3df)* = 35.06, $p < 0.0001$]. This finding reported in [2] offers strong support for the notion that *Personal Predispositions*, which represent relatively stable psychological traits, should have threat ratings that are less likely to decay over time.

3 Characterizing Indicator Decay

3.1 Decay Functions

Let the value of a variable – say the "threat score" – be denoted ***S(t),*** where ***t*** represents time. We wish to characterize the value of ***S*** at time ***t***, given it was observed at time t_0 where $t_0 < t$. We may define the original value of ***S*** (at time t_0), as S_0.

Constant decay function: A constant decay function assumes that the amount of decay is a constant (δ) proportion of the initial value ($0 \leq \delta \leq 1$), i.e., the decrement is $S_0 \delta$ per each time epoch, and the decay function may be written as:

$$S(t) = S_0 (1 - t\delta) \quad (1)$$

A drawback of using this constant decay function is that eventually the value will go negative, which is undefined.

Exponential decay function: A general and extensively used mathematical model for decay is the exponential decay function. In exponential decay, the amount of decay is proportional to the original value of the variable. The exponential decay function may be expressed as:

$$S(t) = S_0 e^{-\alpha t} \quad (2)$$

where e = Euler's constant and α is a constant that determines the rate (percentage) of decay. We can derive an expression for the time it takes for the substance to exponentially decay to a specific value, S_1 by solving equation (2) for t:

$$t = - \ln (S_1/S_0) / \alpha \quad (3)$$

and the "half-life" of a variable—the time it takes for a substance to exponentially decay to *half* of its original quantity ($S_1/S_0 = 0.5$)—is computed as $t_{half\text{-}life} = - \ln (.5)/\alpha = 0.693147/\alpha$.

We can solve for the decay constant α, instead of solving for t, to derive an estimate of the decay constant based on the half-life, i.e., $\alpha = 0.693147/t$. We can get an "intuitive" idea about the value of α for a particular example by asking ourselves how long we think it would take a risk score to decrease to half its value and use this simple equation to find α. For example, if we imagine that it will take 6 months (~180 days) for a threat/risk score to reach half its original value, then we would find $\alpha = 0.693147/180 = 0.00385$.

Another form of the exponential decay function is: $S(t) = S_0 (1-\delta)^t$ where the parameter δ represents the decay constant. This is equivalent to the exponential decay function $S(t) = S_0 e^{-\alpha t}$ since the term $e^{-\alpha}$ corresponds to the term $(1-\delta)$. That is, we can equate the decay parameters, $\delta = 1 - e^{-\alpha}$ or $\alpha = \ln (1-\delta)$.

3.2 Choice of Decay Function

How important is the specific form of the decay function? Table 1 compares the exponential and constant decay rates for a hypothetical case either using an exponential decay or a constant decay, with an assumed initial value of 100 (therefore, one can more generally interpret the figures as percentages of the initial value). For purposes of illustration, consider a decay parameter that corresponds to a decrement of about 8% per month (as reported in [2]). As noted above, this monthly decrement corresponds to a daily constant decay rate, with parameter $\delta = 0.00267$. For an exponential decay model, this translates to $a = \ln (1-0.00267) = 0.00237$. As shown in Table 1, there is little practical difference in the values over the first month and perhaps even extending to several months, considering the low precision of relevant studies in estimating decay parameters (in particular, the linear decrease observed in [2] over 3-4 months is not easily distinguished from an exponential rate). Thus, it may not be feasible, at this time, to decide whether one decay function is more accurate than another. We shall assume that the more general exponential decay model applies when characterizing PRI decay.

Table 1: Constant Versus Exponential Decay Functions

Elapsed Time (t) in Days	Constant Decay (δ=0.00267) Score at time (t)	Exponential Decay (α = 0.00237) Score at time (t)
0	100 = 100 * (1-0*0.00267)	$100=100*e^{-(.00237)*0}$
30	92 = 100 * (1-30*0.00267)	$93=100*e^{-(.00237)*30}$
60	84 = 100 * (1-60*0.00267)	$87=100*e^{-(.00237)*60}$
90	76 = 100 * (1-90*0.00267)	$81=100*e^{-(.00237)*90}$
120	68= 100 * (1-120*0.00267)	$75= 100*e^{-(.00237)*120}$
150	60= 100 * (1-150*0.00267)	$70= 100*e^{-(.00237)*150}$

Figure 1 provides illustrations of five levels of decay ranging from Very High to Very Low (decay parameters are: Very High, $\alpha = 0.15$; High, $\alpha = 0.025$; Medium, $\alpha = 0.012$; Low, $\alpha = 0.00385$; Very Low, $\alpha = 0.002$).

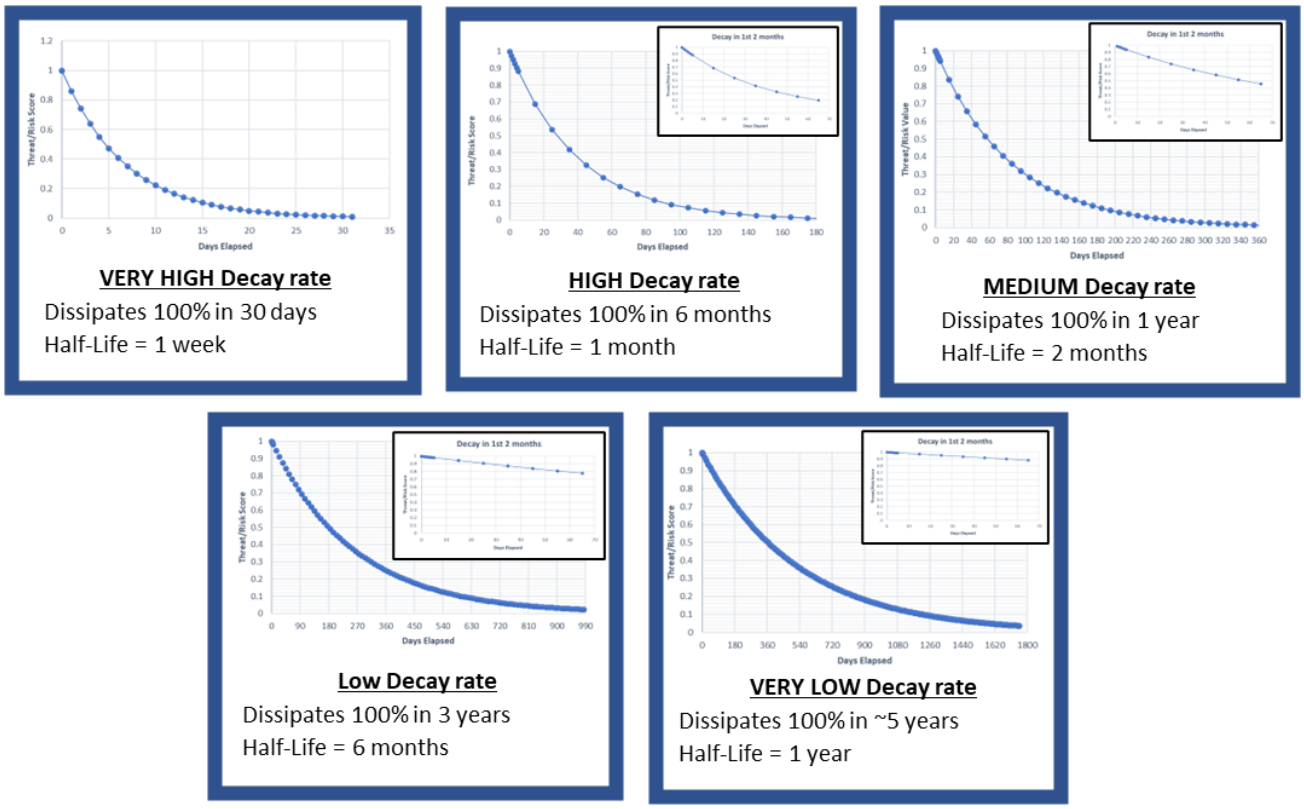

Figure 1. Illustrative Exponential Decay Rates

4 Research Hypotheses

Based on the relatively limited relevant research, we may only speculate about the expected findings when soliciting expert judgments about temporal factors underlying the influence of PRIs on insider threat assessment. We expect that temporal/decay effects should depend upon the nature of the PRI. For cyber/technical indicators that are transient in nature (e.g., login errors), we expect that the influence on insider threat risk assessment will dissipate relatively soon after having been observed or recorded. Nevertheless, some technical PRIs may be of such concern that their influence on threat assessment will not dissipate very much over time: examples are PRIs like installing a password cracker or creating a "backdoor"—reflecting severe policy, legal, or security violations. Similarly, behavioral PRIs such as performance issues or absenteeism may have transient impacts on risk judgments, but psychological factors such as personality traits – which tend to be stable over time [15]—might be expected to have a more durable impact with little or no decay. Thus, it is possible that we may find more than one representative decay rate within a Role Type category, as described in the following research hypotheses:

1. **High and Medium decay rates for Technical Precursors**. Generally, Technical Precursors will have higher decay rates than other types of PRIs; however, extremely serious Technical Precursors, such installing a keylogger, will have longer lasting effects on threat assessment judgments than more transient technical PRIs.
2. **Low to No-decay rates for Personal Predispositions** Personal Predispositions that represent psychological factors and personality traits are expected to have longer-lasting impacts on threat assessment judgments and should decay relatively slowly, if at all.
3. **Medium to Low decay rates for Behavioral Precursors**. Behavioral PRIs, like Technical Precursors, may come in two varieties—those that have a relatively transient impact and those that have a more persistent impact on threat judgments. Occasional concerning behaviors that reflect more transient concerns (such as performance issues or absenteeism) might demonstrate short-lived impacts on threat judgments (moderate decay rate), but

some that are more egregious (such as workplace violence, weapon misuse) may demand a longer-lasting influence and yield low rates of decay.

4. **Medium decay rates for Precipitating Events**. There is a paucity of relevant research on which to base expectations. Precipitating events might well be time-constrained (e.g., pending relocation, transfer, layoff, retirement, etc.); some that represent past stressors (poor performance ratings, possible reduction in force, etc.) may produce malicious motivations that could be long-lasting.

5 Method

5.1 Materials

To study decay characteristics of PRIs, we merged the indicators defined in SOFIT and the DoD taxonomy, producing a set of ~265 PRIs. Because we're focusing on decay characteristics, it seemed most appropriate to follow [2] in categorizing the PRIs along the dimension referred to as "Role Type" since preliminary indications are that PRIs might differ in their decay parameters based on Role Type. Four main categories of Role Type are:

- ***Precipitating Event.*** An event that triggers or motivates the insider to carry out an insider crime. [Examples: *disciplinary action, passed over for promotion, revocation of security clearance*]
- ***Personal Predisposition.*** A (personal) characteristic historically linked to a propensity to exhibit malicious insider behavior. [Examples: *gambling addiction, mental instability, self-harm, suicidal ideation*]
- ***Behavioral Precursor.*** An individual action, event, or condition that involves personal or interpersonal behaviors and that precedes and is associated with insider activity. [Examples: *attempts to obtain national security information without need-to-know, criminal behavior involving weapons, verbal abuse/bullying*]
- ***Technical Precursor.*** An individual action, event, or condition that involves computer or electronic media and that precedes and is associated with malicious insider activity. [Examples: *disabling anti-virus software*, *excessive use of screen capture*, *sending E-mail to suspicious address*]

5.2 Procedure

A knowledge elicitation survey was conducted with twelve experts at a DoD insider threat hub. The survey, which was implemented in Microsoft Excel, included an introductory tab explaining the purpose of the study and an Informed Consent page; a tab with a short list of demographic questions; and a tab with a list of the 265 PRIs, organized into sections corresponding to the four PRI Role Types. The instructions are provided in Appendix A.

The first paragraph of the instructions included a brief description of the possible nature of the PRI decay construct to raise the consciousness of participants about possible differences in decay rates. This contrasts with the procedure used in [2] that did not provide such context (but still found a difference between decay probabilities for Personal Predispositions versus other PRI types). While our procedure may produce demand characteristics that could bias responses, we wished to explain and illustrate the concepts that were being studied in the survey, to encourage due consideration of possible decay rate differences.

A response was required for each PRI, with six decay-rate response options: Very High, High, Medium, Low, Very Low, No-Decay/None. The task was straightforward and only required relatively quick judgment; we estimated that approximately one hour was required to complete the task (if done in one continuous session), but the participant could save work and exit/return at any time. For this study, we enlisted volunteers from a Department of the Air Force laboratory who were able to complete the surveys within one week. Some of the demographic data were garbled or lost, but a general description of the participants is as follows. Six of the twelve participants were in the 31-45 age range; six were in the 46-60 age range. Nine of the twelve described themselves as Practitioners; three identified as Researchers. Five participants identified counterintelligence or insider threat as their primary discipline; two identified as computer scientists; the remaining disciplines cited included sociology, political science, psychology, and criminal justice.

6 Results

The interrater reliability, measured by Fleiss' Kappa, for each of the four Role Types separately, with resulting "fair to moderate" effect-size Kappa measures of 0.50, 0.32, 0.44, and 0.36 for Precipitating Events, Predispositions, Behavioral Precursors, and Technical Precursors (respectively). The 6x4 contingency table was analyzed using a chi-square test to reveal a highly significant association ($\chi^2 = 194.9$ with 15 *df*) between decay rate and Role Type. The Cramer's V effect-size

statistic is V = 0.14, which approaches a medium effect size [16].

To ease the interpretation of data, we combined the six decay response categories into four decay rate categories (None, Low/Very Low; Medium; and High/Very High); the distribution of PRIs assigned to these decay rates is shown for each of the four Role Types in Figure 2 (the calculated chi-square values and effect sizes are similar for this contingency table). The distribution of decay rates by Role Type is further illustrated in Table 2, which provides a list of PRIs that were *most consistently* assigned (by at least six of the twelve experts) to one of the four decay rate categories. The blank cells in the last column of the table, representing the High/Very High decay rate category, reflects the relatively low preference for the highest decay rates as well as somewhat low agreement across expert judges.

Examination of Figure 2 and Table 2 reveals distinct differences across and within Role Types:

- Technical Precursors are much less likely to be assigned a no-decay rating (only 9% of these indicators were rated in the no-decay category). Over one-half (56%) of the Technical Precursors were assigned Medium to Very High decay rates. This finding is consistent with our conjecture, in Hypothesis 1, that these indicators would likely be given medium to high decay rates.

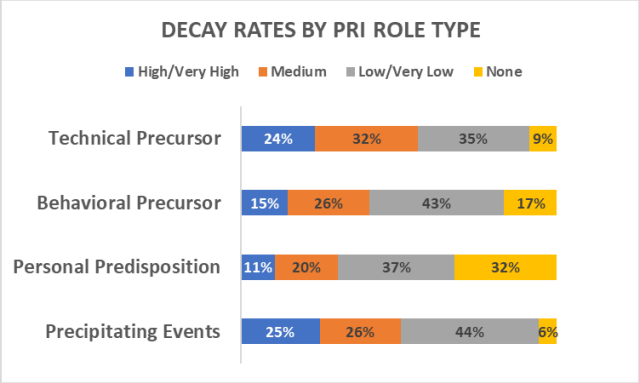

Figure 2. Distribution of Decay Rates by PRI Role Type

- Personal Predispositions, in contrast, were most likely to receive no-decay or low decay rate estimates (69%). They were least likely to be identified with high or very high rates of decay (11%). This association of Personal Predispositions with no-decay or low decay rates is consistent with Hypothesis 2 and agrees strongly with previous results reported in [2], which used a very different methodology.

Table 2: Most-Consistently-Assigned PRI Decay Rates by Role Type [a]

Role Type	Decay Rate Characterization			
	No Decay	**Very Low/Low**	**Medium**	**High/Very High**
Precipitating Event	•	• Security clearance suspension	• Reprimand	• Pending transfer or relocation
Personal Predisposition	• Insanity plea • Psychopathy • Narcissism • Grandiosity • Lack of remorse	• Involuntary treatment for substance abuse	• Negative characterization of previous employment	•
Behavioral Precursor	• Associating with extremist or terrorist groups • Workplace violence • Advocating terrorism or violence	• Delinquent debts • Substance abuse • Travel policy violation	• Adverse changes to financial status • Declining performance • Security violation • Attendance issues	•
Technical Precursor	• Introduction of malicious code	• Unauthorized storage device • Unauthorized wireless	• Large data transfers • Change file extensions • Printing to anomalous location	•
Tentative/Recommended Decay rate range	$\alpha \sim 0.0$	$0.00385 < \alpha < 0.002$	$\alpha \sim 0.012$	$0.025 < \alpha < 0.15$

[a] PRIs shown here received relatively consistent decay ratings across expert participants (assigned the same decay rate by at least six of the twelve experts).

- Behavioral Precursors were very likely to be considered to have no-decay or low decay rates (70%). Forty-one percent of the Behavioral Precursors were assigned to the Medium-to-Very High decay rate categories. This is consistent with Hypothesis 3.
- Precipitating Events were very unlikely to be assigned a no-decay rating (6%) but were relatively likely to be considered to have very low to medium decay rates (70%), which is consistent with our conjecture in Hypothesis 4.

While not depicted in the figure (which shows combined ratings in the High and Very High decay rate categories), we observe in our full data set that our expert analysts seemed reluctant to assign the Very High decay rate category to PRIs—only 4% of Behavioral Precursors and Personal Predispositions were assigned the highest decay rate; 6% of Precipitating Events were assigned the Very High decay rate; and 8% of Technical Precursors were assigned this rate of decay. This may reflect the fact that the analysts wish to avoid overlooking issues of concern. Even more telling (as seen in the paucity of consistently rated PRIs listed in the last column of Table 2) is that there was low agreement across raters in assigning PRIs to the highest decay rate.

Notwithstanding possible minor differences between our expectations and these results, these findings confirm the general observation that insider threat indicators decay at different rates, with some distinctive differences based on indicator Role Types. The last row of Table 2 suggests tentative decay rate ranges that may be applied to PRIs based on the current findings. Since there are variations within Role Types, it may be advisable to break down the Role Type classes into two or three subclasses for purposes of assigning decay rates, following the guidance in this table.

7 Discussion

We conducted an expert knowledge survey that provided minimal background with brief examples to "educate" participants about possible considerations that may affect insider threat indicator decay characteristics. While these instructions could produce demand characteristics that potentially bias the judgments, we deemed it useful to call attention to the possible nature of the decay construct (in contrast to the procedure employed in [2], which provided no such context). Despite this, interrater reliability was moderate at best, but the results were nevertheless statistically significant, and we were able to infer statistically reliable differences in judged decay rates by Role Type.

Challenges remain to devise methods that increase interrater reliability when expert analysts with diverse backgrounds are used. For example, a meeting or training session focused on raising understanding/awareness of temporal factors could be conducted prior to the survey study. Hubbard [17] notes that even a small amount of training (e.g., 3 hours) can significantly improve calibration of estimates. Alternatively, a structured group-consensus approach, such as a Delphi study [18], may be used to estimate decay rates. A third method—used in [2]—derived decay rate estimates based on judgments of level of risk for indicators and combinations of indicators that occur at different times along a scenario timeline.

8 Conclusions

The results of this study revealed challenges—which are by no means unexpected (e.g., see [19]—in obtaining reliable human judgments of parameters needed to build and test models of human decision making and judgment. Our results are nevertheless useful in informing future studies. While the PRI decay properties reported in [2] manifested in expert judgments of threat severity, the present study sought to acquire judgments of decay characteristics directly by obtaining expert judgments of the time span of PRI influence on judgments of insider threat risk. Plans for a larger or more formal study may seek to obtain judgments using both methods.

While this study adds to the small body of research on PRI decay, other interesting research topics remain unexamined, such as the possible nonlinear effects that may occur when a past (possibly decayed) PRI is observed again. For example, suppose a security infraction (with a medium decay rate) occurred one year ago, so that its risk score will now be near zero. How will another security violation impact the judged insider risk? A simple model may treat these as independent events, while a more complex model would recognize and account for such temporal interaction patterns. These effects can be studied using methods such as the survey conducted in [2], which asks for judgments of risk for combinations of PRIs while also varying temporal factors. It would not be possible to assess such effects in the present study, but it is interesting to note that some of our expert analyst participants raised similar questions about possible nonlinearities and PRI interactions. For example, unsolicited analyst feedback included:

- "Will interaction with another PRI increase or decrease the decay? Should slower decay (rate) be kept when interactions occur?"

- "There is an interaction among these PRIs (IMHO)…. If I see someone using a Priv Account, then Deleting/Modifying Audit Logs, on a regular basis, … my "care factor" is very low (Very High decay rate) as I surmise this is likely their job to perform. But someone doing/attempting the same action but without a Privileged Account … is much more concerning and remains on my radar for a longer period of time."

The second example not only speaks to the need for further research to account for temporal effects and PRI interactions, but also to weigh in the influence of other factors, including the person's job role.

We may offer preliminary, and speculative, conclusions based on our findings. The results reported here, using a substantially different methodology and a far more comprehensive set of PRIs, generally agree with findings reported in [2]—which increases the face validity of the results. However, the current findings also suggest that the Role Type classification alone may be insufficient for classifying decay rates. As a practical matter, based on the present findings shown in Figure 2—which should be considered tentative until replicated or validated—we may suggest a simple, expedient assignment of exponential decay parameters based on Role Type: Precipitating Events generally assigned medium decay, $\alpha \sim 0.01$; Personal Predispositions generally assigned no decay, $\alpha \sim 0$; Behavioral Precursors generally assigned low/very low decay, $\alpha \sim 0.003$; and Technical Precursors generally assigned relatively high rate of decay, $\alpha \sim 0.02$. A more nuanced assignment should follow the guidance reflected in the ranges of exponential decay rates that vary both across and within Role Types, as shown in the last row of Table 2.

The main contribution of this research is the discussion of applying indicator decay in assessing threat, the description of its implementation using exponential decay, and the rationale for using indicator Role Type as an initial approach to distinguish indicator decay rates. Once developed and validated, it is straightforward to integrate an indicator decay "operator" into insider threat assessment models such as those studied in [7] and [8], among others. In research and development that we are conducting for the Department of the Air Force, we shall be incorporating and evaluating these decay concepts in a continuous intelligence platform, called Cogynt, which uses a hierarchical, complex event processing approach that is well-suited to the type of pattern-based analyses that we believe will enhance the level of decision support for insider threat analysts.

ACKNOWLEDGMENTS

This research was supported by Contract FA7146-20-C-0955 conducted for the United States Department of the Air Force Counter-Insider Threat Program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the U.S. Government. The authors gratefully acknowledge useful suggestions by peer reviewers.

REFERENCES

[1] CERT Insider Threat Center. 2016. Common Sense Guide to Mitigating Insider Threats, 5th ed., Carnegie Mellon University Software Engineering Institute, Technical Note CMU/SEI-2015-TR-010. https://resources.sei.cmu.edu/asset_files/TechnicalReport/2016_005_001_484758.pdf. [online: accessed on September 30, 2022]

[2] Frank L. Greitzer and Justin Purl. 2022. The dynamic nature of insider threat indicators. Springer Nature Computer Science, 3(102). https://doi.org/10.1007/s42979-021-00990-1. [online: accessed on September 30, 2022]

[3] George B. Magklaras and Steven M. Furnell. 2005. A preliminary model of end user sophistication for insider threat prediction in IT systems. *Computers & Security,* 24(5), 371-380. https://doi:10.1016/j.cose.2004.10.003 https://www.sciencedirect.com/science/article/pii/S0167404804002603 [online: accessed on September 30, 2022]

[4] Philip A. Legg, Nick Moffat, Jason R. C. Nurse, Jassim Happa, Ioannis Agrafiotis, Michael Goldsmith, and Sadie Creese. 2013. Towards a conceptual model and reasoning structure for insider threat detection. *Journal of Wireless Mobile Networks*, *Ubiquitous Computing, and Dependable Applications*, 4(4), 20-37. https://doi.org/10.22667/JOWUA.2013.12.31.020 [online: accessed on September 30, 2022]

[5] Jason R. C. Nurse, Oliver Buckley, Philip A. Legg, Michael Goldsmith, Sadie Creese, Gordon R. T. Wright, and Monica Whitty. 2014. Understanding insider threat: A framework for characterising attacks. *IEEE Security and Privacy Workshops (SPW)*, San Jose, CA (pp. 214-228). IEEE. http://ieeexplore.ieee.org/document/6957307/?arnumber=6957307 [online: accessed on September 30, 2022]

[6] Frank L. Greitzer, Lars J. Kangas, Christine F. Noonan, Chris R. Brown, and Thomas Ferryman. 2014. Psychosocial modeling of insider threat risk based on behavioral and word use analysis. *e-Service Journal*, 9(1), 106-138. http://www.jstor.org/stable/10.2979/eservicej.9.1.106 [online: accessed on September 30, 2022]

[7] Frank L. Greitzer, Justin Purl, Yung-Mei Leong, and D. E. (Sunny) Becker. 2018. SOFIT: Sociotechnical and Organizational Factors for Insider Threat. In *2018 IEEE Security and Privacy Workshops*, San Francisco, CA, May 24, 2018.

[8] Frank L. Greitzer, Justin Purl, D. E. Becker, Paul J. Sticha, and Yung-Mei Leong. 2019. Modeling expert judgments of insider threat using ontology structure: Effects of individual indicator threat value and class membership. *52nd Hawaii International Conference on Systems Sciences (HICSS-52)*, Maui, Hawaii, January 2019, pp. 3202-3211.

[9] Daniel L. Costa, Matthew L. Collins, Samuel J. Perl, Michael J. Albrethsen, George J. Silowash, and Derrick L. Spooner. 2014. An Ontology for Insider Threat Indicators. In K. B. Laskey, I. Emmons and P C.G. Costa (Eds.), *Proceedings of the Ninth Conference on Semantic Technologies for Intelligence, Defense, and Security* (STIDS 2014), 2014, 48–53. https://apps.dtic.mil/sti/citations/ADA615757 [online: accessed on September 30, 2022]

[10] Frank L. Greitzer, Justin Purl, Paul J. Sticha, Martin C. Yu, and James Lee. (2021). Use of Expert Judgments to Inform Bayesian Models of Insider Threat Risk. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 12(2), 3-47. June 2021*. DOI:10.22667/JOWUA.2021.06.30.003. https://dx.doi.org/10.22667/JOWUA.2021.06.30.003. Supplementary files available at: http://isyou.info/jowua/abstracts/jowua-v12n2-1.htm

[11] H. Keith Hunt, H. David Hunt, and Tracy C. Hunt. 1988. Consumer grudge holding. Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, 1, 116-118.

[12] Sweta C. Thota and Newell D Wright. 2006. Do consumers hold grudges and practice avoidance forever? A Markov Chain model of the decay of grudgeholding and avoidance attitudes. *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, *19*, 89-102. https://www.jcsdcb.com/index.php/JCSDCB/article/download/10/217

[13] Yany Gregoire, Thomas M. Tripp, and Renaud Legoux. 2009. When Customer Love Turns into Lasting Hate: The Effects of Relationship Strength and Time on Customer Revenge and Avoidance. *Journal of Marketing*, *73(6)*, 18-32.

[14] Thomas M. Tripp and Yany Gregoire. 2011. When unhappy customers strike back on the Internet. *MIT Sloan Management Review*, Reprint 52303, pp. 1-8. Massachusetts Institute of Technology. https://chaireomerdesserres.hec.ca/wp-content/uploads/2019/06/Sloan_Unhappy-Customers_Strike-Back_Internet.pdf

[15] Deborah A. Cobb-Clark and Stefanie Schurer. 2012. The stability of big-five personality traits. *Economics Letters*, 115(1), 11-15. https://www.econstor.eu/bitstream/10419/55111/1/675938643.pdf [online: accessed on September 30, 2022]

[16] Jacob Cohen. 1988. Statistical power analysis for the behavioral sciences (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

[17] Douglas W. Hubbard and Richard Seiersen. 2016. *How to Measure Anything in Cybersecurity Risk.* Hoboken, New Jersey: John Wiley.

[18] Harold A. Linstone and Murray Turoff. (eds). 1975. *The Delphi Method: Techniques and Applications*. Reading, MA: Addison-Wesley.

[19] Amos Tversky and Daniel Kahneman. 1974. Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1130.

Appendix A: Instructions Provided to Participants in Expert Knowledge Acquisition Study

The following instructions—and only these instructions—were provided in our expert knowledge acquisition study focused on PRI decay:

> Research and common sense suggest that in addition to the "severity" of any insider threat indicator, the passage of time since observing the indicator must be considered when assessing insider threat risk. For example, a very recent observation of the occurrence of login errors may influence threat assessment, but the same type of incident occurring months or years in the past would not have much impact on judged threat. On the other hand, the impact of psychological factors such as personality traits that are stable over decades should influence insider threat judgments over a very long period. Other types of indicators—such as disciplinary actions or various behavioral factors—may influence risk judgments for intermediate time periods (e.g., weeks, months).
> We refer to this temporal factor as Indicator Decay and we wish to estimate decay rates. Generally, an indicator's influence gradually decreases based on its rate of decay (its half-life). We distinguish six possible decay rates:
>
> - VERY HIGH—the impact of an indicator is transient: its influence is reduced by 50% each week ("half-life" = 1 week) and it influences risk judgments only for about a month.
> - HIGH—the indicator influences risk judgments only for about 6 months; half-life = 1 month.
> - MEDIUM—reflects a 1-year timeframe with a half-life of 2 months.
> - LOW—corresponds to a 3-year period during which the indicator may influence threat judgments; half-life = 6 months.
> - VERY LOW—the indicator's influence continues to impact threat judgments for 5 years; its influence is reduced by 50% per year (i.e., its half-life is 1 year).
> - NO-DECAY—the indicator maintains its original threat value; an analyst will always take the indicator into account regardless of time elapsed since its occurrence.
>
> This exercise examines the influence of time on judgments of insider threat for different types of indicators. Because we anticipate that insider threat indicators might have different rates of decay – how long they influence expert judgments of risk – based on the types of indicators being considered, we group insider threat indicators into four "role-type" categories:
>
> - ***Precipitating Event.*** An event that triggers or motivates the insider to carry out an insider crime. [Examples: *disciplinary action, passed over for promotion, revocation of security clearance*]
> - ***Personal Predisposition.*** A (personal) characteristic historically linked to a propensity to exhibit malicious insider behavior. [Examples: *gambling addiction, mental instability, self-harm, suicidal ideation*]
> - ***Behavioral Precursor.*** An individual action, event, or condition that involves personal or interpersonal behaviors and that precedes and is associated with insider activity. [Examples: *attempts to obtain national security information without need-to-know, criminal behavior involving weapons, verbal abuse/bullying*]
> - ***Technical Precursor.*** An individual action, event, or condition that involves computer or electronic media and that precedes and is associated with malicious insider activity. [Examples: *disabling anti-virus software, excessive use of screen capture, sending E-mail to suspicious address*]
>
> In this survey, we will ask you to indicate your judgment of indicator decay rate for a large number of insider threat indicators, organized by role type. First, we will ask for some brief demographic information about you. Next, we will provide the list of insider threat indicators and ask that you make a decay-rate "selection" (one of the 6 rates defined above) for each indicator. Please save this sheet often so you will not lose information as you complete the task. You needn't complete all items in one sitting, but please do try and provide consistent ratings as you go through the exercise.