21
10-23-80
2 to NTIS

SAND80-0445
Unlimited Release

PERFORMANCE ESTIMATES FOR PERSONNEL ACCESS CONTROL SYSTEMS

R. G. Bradley

Sandia National Laboratories

SAND80-0445

Unlimited Release
Printed August 1980

PERFORMANCE ESTIMATES FOR PERSONNEL ACCESS CONTROL SYSTEMS

R. G. Bradley
Sandia National Laboratories
Entry Control Systems Division 1727
Albuquerque, New Mexico 87185

ABSTRACT

Current performance estimates for personnel access control systems use estimates of Type I and Type II verification errors. A system performance equation which addresses normal operation, the insider, and outside adversary attack is developed. Examination of this equation reveals the inadequacy of classical Type I and II error evaluations which require detailed knowledge of the adversary threat scenario for each specific installation. Consequently, new performance measures which are consistent with the performance equation and independent of the threat are developed as an aid in selecting personnel access control systems.

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PERFORMANCE ESTIMATES FOR PERSONNEL ACCESS CONTROL SYSTEMS

Summary

One essential part of a total safeguards concept to protect a facility is the personnel access control system. Ideally, such a system should only allow normal movement of authorized personnel to their usual workplace while at a minimum, detecting and delaying unauthorized entry or exit from a protected area. Unfortunately, no system whether manual, machine-aided manual, or automated, has been found to be invulnerable to errors. The safeguard systems designer in selecting an access control system for a specific facility must use some set of performance measures to select the "best" system available. Obviously, the level of security and the relative ease of use of a personnel access control system are of primary importance in selecting a system for a particular facility.

In the past, classical statistical estimates of the Type I error (false rejection of an authorized individual) and Type II error (false acceptance of an imposter) have played a primary role in estimating the performance of a personnel access control system. However, because of the difficulty and confusion in quantifying these error rates, comparisons between access control systems have not always been consistent. Consequently, a performance equation that encompasses, not only normal entry, but the insider and outside adversary threat has been developed. Access attempts are categorized by major attributes which include proper identification as well as possible acts of tampering. The classical Type I and Type II error estimates, when defined in context of this performance equation, prove to be highly dependent on the adversary threat as well as the particular personnel access control system employed.

Examining the performance equation for a facility reveals that the likelihood of gaining access to it can be partitioned into six factors relating to normal use, insider threat, and outside adversary threat. Each of these terms can again be divided into two components, one dependent on the population and/or the adversary threat and one which is characteristic cf the access control system used. Three new performance measures are presented to replace the inadequate Type I and II errors.

The first is a measure of consistency. Essentially the complement of the Type I error, this measure is an estimate of the relative ease of use and user acceptance that a particular access control system affords. This consistency measure, which is easily determined from the performance data, can also be used to predict the number of bypass or exception cases that must be allowed for by the facility designer.

The second measure is the uniqueness that a particular access control algorithm attains. By determining the likelihood of random identity crossing in the enrolled user population, this measure of uniqueness or separability can be used by the facility designer to determine the relative security that a particular system affords against a threat in which no active methods of deceit are employed.

The third measure is the resistance to tampering. Though this measure can never be completely quantified like the other two, it probably is the most important for a facility in which the adversary threat is severe.

Finally, guidelines on how to determine and use these three performance factors -- consistency, separability, and resistance to tampering -- are developed to aid the facility designer in selecting a personnel access control system.

Background

One essential part of a total safeguards concept to protect a facility is the personnel access control system. Ideally, such a system should only allow access to a specific area, material, or action by those people who have been authorized. Unfortunately, no system, whether manual, machine-aided manual, or automated, has been found to be invulnerable to errors. The safeguard systems designer, in selecting an access control system for a specific facility, should employ some performance measure in order to select the "best" system available. Present techniques to evaluate and compare different access control systems, though similar in intent, do not always use the same method to estimate performance. The following discussion is intended to establish some mathematical rigor and uniformity for the evaluation of access control systems for possible implementation in a total safequard system.

An access control system deals with people; consequently, user acceptance, ease of use, and a whole gamut of human factors must be considered along with initial and operational costs. Thus, the process of selecting a system deals with the tradeoffs between security and the costs, both human and otherwise, of the system. To quantify both the security and relative ease of use of each system being considered, some standardized performance measure must be applied. To establish this performance equation, a probabilistic approach will be taken necessitating the following assumptions:

1. Each encounter with the access control system will be considered an access attempt. Access attempts can be categorized into two groups based on outcome, success (S) or failure (F). The sum or union of these mutually exclusive results constitutes the universe or set of all possible access attempts (Ω). Likewise, the intersection of these two mutually exclusive sets is the null or empty set (ϕ). These are denoted by

   $S+F = \Omega$, and

   $S \cdot F = \phi$.

2. Access attempts can also be partitioned into attempts of any method that were performed by an enrolled individual (E) or a person not enrolled

who is outside the system or otherwise unauthorized (U). Again, these two mutually exclusive sets can be denoted by

$E+U = \Omega$, and

$E \cdot U = \phi$.

3. Finally, the set of access attempts can be partitioned into three mutually exclusive sets based on the method of the access attempt. First, an individual can use the system in a normal manner and claim that he is himself (G). Second, an individual could use the system in a normal manner but claim to be someone else, either by accident or purposely (B). Third, a person can use the system in an abnormal manner, tamper, counterfeit, or otherwise attempt to actively defeat the system (T). These three subsets conform to the following notation.

$G+B+T = \Omega$

$G \cdot B = \phi$

$G \cdot T = \phi$

$B \cdot T = \phi$.

The following notation will be used.

$P(X)$ = The probability that an event from the set X will occur.

$P(X/Y)$ = The conditional probability that an event from the set X will occur given that an event from the set Y has occurred.

Performance Equation

For the moment we will consider only the subset of access attempts in which the outcome was a success. Because in each case the set of all access attempts was partitioned into mutually exclusive subsets, the subset of successful access attempts can itself be partitioned into mutually exclusive subsets which can be represented by

$S \cdot E \cdot G$

$S \cdot E \cdot B$

$S \cdot E \cdot T$

$S \cdot U \cdot G$

$S \cdot U \cdot B$

$S \cdot U \cdot T$

In addition, the probability of a successful access attempt becomes the sum of the probabilities of the members of the subsets, or

$$\begin{aligned} P(S) = {} & P(S \cdot E \cdot G) + P(S \cdot E \cdot B) + P(S \cdot E \cdot T) + \\ & P(S \cdot U \cdot G) + P(S \cdot U \cdot B) + P(S \cdot U \cdot T) \ . \end{aligned} \qquad (1)$$

From the law of compound probability[1]

$$P(X \cdot Y) = P(X/Y)P(Y) \ ,$$

the probability of a successful access attempt (which will be considered the performance equation for an access control system) can be written as

$$\begin{aligned} P(S) = {} & P(S/E \cdot G)P(E \cdot G) + P(S/E \cdot B)P(E \cdot B) + \\ & P(S/E \cdot T)P(E \cdot T) + P(S/U \cdot G)P(U \cdot G) + \\ & P(S/U \cdot B)P(U \cdot B) + P(S/U \cdot T)P(U \cdot T) . \end{aligned} \qquad (2)$$

For completeness, the probability of failure can likewise be written

$$\begin{aligned} P(F) = {} & P(F/E\cdot G)P(E\cdot G) + P(F/E\cdot G)P(E\cdot B) + \\ & P(F/E\cdot T)P(E\cdot T) + P(F/U\cdot G)P(U\cdot G) + \\ & P(F/U\cdot B)P(U\cdot B) + P(F/U\cdot T)P(U\cdot T) \end{aligned} \tag{3}$$

and combining Eqs. (2) and (3), we note that

$$P(S) + P(F) = 1. \tag{4}$$

Examining the performance, Eq. (2), we can make some general observations. Although the performance of an access control system cannot be separated from the facility and population for which it has been installed, the 12 factors that make up the 6 terms in the performance equation can be separated into two groups, one dependent on the particular access control system and the other a function of the user population. The conditional probabilities are the response of the access control system to a specific user population and tend to be independent of the facility while the joint probabilities are the representation of the user population and/or threat to the facility itself.

The six joint probabilities represent the likelihood that a specific access attempt will occur regardless of outcome. The access attempts can also be classified as to what member of the user population the access attempt comes from. The first subset, E·G, is the employee of the facility using the system in a normal manner in the course of his job. The next two subsets, E·B and E·T, form the insider threat and an ideal access control system would be able to recognize this and deny access so as to prevent unauthorized and illicit acts. It is admitted that an insider could also sabotage a facility from his area of authorized access but the access control system can neither recognize nor prevent this action. Finally, the subsets U·G, U·B, and U·T constitute the outside adversary threat and must also be denied access. Thus, the access attempts can be classified as

E·G = The normal employee (Category I)

E·(B+T) = The insider threat (Category II)

U·(G+B+T) = U = The outside adversary threat (Category III)

Performance Estimates

In access control it is common to consider two kinds of verification errors for a personnel identification system.[2-7]

Type I - The rejection of an enrolled individual who purports to be himself (Category I individual).

Type II - The acceptance of an unauthorized individual or an enrolled individual acting in an abnormal manner (Categories II and III individuals).

The errors are an adaptation to the Type : and Type II errors used in classical statistics and are perfectly valid concepts for an access control system. However, the 12 terms that generate the performance Eq. (2) also are necessary to evaluate these errors. Thus, for an operational access control system at a particular facility, the user population and the actual threat (which is likely to be different from the design threat) will affect the quantification of these verification errors.

To illustrate effects, assume that we have the same access control system at two identical facilities with the exception that one facility is protecting somethi.ıg perceived by everyone as worthless and uninteresting while the other system is protecting something of extreme value. Why would anyone desire access to the first facility unless it was a job requirement? Thus, the likelihood of a threat to a facility depends on what is being protected and the "value" of the unauthorized action. Furthermore, the magnitude of the penalty imposed for an unsuccessful unauthorized access attempt will also affect the type of illicit attempt. If the penalty is merely a slap on the wrist and an empty warning not to try it aqain, why would an adversary use extraordinary means if pure chance and repeated trials would succeed? However, if the penalty was death, it would be reasonable to assume that an intruder would try to find a method to maximize his likelihood of success. Thus, the more severe the penalty, the higher likelihcod that an adversary would tamper with or otherwise attempt to defeat the system rather than casually chance being caught using the system in a normal manner.

The effect of the magnitude and makeup of the threat on the error rates is easily seen when the veri..ication errors are derived from the performance equation and compared to the performance data available. The verification errors are

$$\text{Type I} = 1 - P(S/I) \qquad (5)$$

$$\text{Type II} = \frac{P(S/II)\ P(II) + P(S/III)\ P(III)}{P(II) + P(III)} \qquad (6)$$

where

$$P(I) + P(II) + P\ (III) = 1$$

$$I = E\cdot G, \text{ the normal user}$$

$$II = E\cdot(B+T), \text{ the insider}$$

$$III = U, \text{ the outside adversary}$$

Short of interviewing each person as he/she uses the access control system (which would bias the results), the only performance data generally available to evaluate the access control system at an actual facility are:

1. The total number of access attempts, $|\Omega|$

2. The number of successes, $|S|$, and failures, $|F|$, but not the category (I, II, or III) the attempt came from

3. The number of reference files and characteristics of the enrolled individuals.

Obviously, this is not enough information to evaluate the verification error rates. Because these verification errors are so dependent on the magnitude of the threat to the facility (P(II), P(III)), Type I and Type II verification errors should not be used by the system designer as a basis for selecting an access control system. Selection of a system for a particular facility should not be based on the likelihood of a penetration attempt, but on how well the system deals with an adversary once the attempt is made. Thus, the performance estimates needed by the facility designer are the conditional probabilities contained in the performance equation, Eq. (2). These conditional probabilities are the response of the system to a specific type of access attempt and are independent of the threat magnitude required to evaluate the Type I and Type II errors. The conditional probabilities are:

$$P(S/E\cdot G)$$

$$P(S/F\cdot B)$$

$P(S/E \cdot T)$

$P(S/U \cdot G)$

$P(S/U \cdot B)$

$P(S/U \cdot T)$

Unfortunately, three of the terms require knowledge of the outside adversary population and thus cannot be exactly evaluated. However, some reasonable bounds can be put on them.

The skills, familiarity, and physical characteristics of the outside adversary can only be hypothesized, but in all probability are not any better than the dedicated insider who as a user is at least familiar with how the access control system works. Thus, it is reasonable to assume that the P(S/E) is greater for any specific method of access (G, B, or T) than the P(S/U). This is obviously true for $P(S/E \cdot G)$ and $P(S/U \cdot G)$; the latter should be negligible because the access control system does not even know that individual exists. One further simplification is possible. The class of access attempts B includes identity mismatches, both accidental and intentional. However, only the subset B_1, in which that identity is also enrolled and therefore recognized by the access control system, has any significant chance of being granted access. Thus, the necessary performance estimates required to bound those from the performance equation are

$P(S/E \cdot G)$

$P(S/E \cdot B_1)$

$P(S/E \cdot T)$.

The first $P(S/E \cdot G)$ represents the likelihood of an authorized user obtaining access. It is essentially a measure of consistency or the ability of the user to duplicate the reference measurements established during enrollment in the system. In a manual system in which a picture badge is used as the reference set of characteristics, the identity match is based on how well an individual resembles the enrollment photo. Obviously such things like growing a beard, wearing contacts or glasses, and changes in hair style or color can affect the likelihood of an identity match. Thus, the user consistency in matching the reference characteristics is also a measure of the ease of use of the system and is estimated from the performance data as the number of successes divided by the number of attempts:

$$P(S/E \cdot G) = |S|/|\Omega| = 1 - |F|/|\Omega|$$

This estimate is detrimentally affected by the outcomes of unauthorized attempts (Category II, III); however, for most access control systems in which the authorized users are the bulk of the access attempts, $P(E \cdot G) \approx 1$, this estimate is a good approximation of the actual conditional probabilitv.

The second performance measure, $P(S/E \cdot B_1)$, represents the likelihood that two individuals possess the "identical" characteristic used to verify identity. For a manual system employing picture badges, it is a measure of how well the guard can use a photograph to distinguish between individuals. Obviously this is fairly subjective and can only be qualitatively determined by random "attacks" on the system. However, for automated personnel identity verification systems (fingerprint, hand geometry, handwriting, voice, etc.) which make use of a physiological characteristic, a quantitative estimate of separability can be made. These automated systems perform the identity comparison based on some fixed algorithm which matches reference or enrollment characteristics against a new set of measured characteristics at each access attempt. Thus, either by cross comparing the enrollment files or by comparing the enrollment files against access attempt characteristics, a quantitative estimate of the separability or uniqueness for the particular algorithm and identifier can be determined. Techniques used in reliability theory can be applied to determine the validity bounds of these data. Obviously, the larger the sample population used in determining this estimate of separability, the better the estimate for that particular system.

The third conditional probability, $P(S/E \cdot T)$, is a measure of the resistance to tampering and/or other methods of active deceit. Unfortunately, this is a subjective measure as there is no way to determine every method of attacking the system. However, some qualitative bounds can be determined. Access control systems can be "blackhatted" by knowledgeable individuals or users to determine the extent of the weaknesses that the systems possess. Admittedly, these attacks on the systems are incomplete and thereby suspect in that they can never address all possible scenarios; however, they should provide some means of relative comparison between systems. Sandia National Laboratories, in their tests of barrier and intrusion detection systems, has found this technique extremely useful.

Application of these estimates by the system designer requires some care and consideration. It must be remembered that any test or collection of performance data on an existing system is limited. The estimates of consistency and separability on any particular system are just that, estimates. Care must be exercised in extending the numbers produced in one test to another situation. Only after several tests on different sites and populations produce consistent results can generalization to the total population be valid. Thus, the system designer will still be forced to examine the test conditions in making his decisions. This is especially true for the third performance estimate, the resistance to tampering. The fact that "blackhat" attacks on one system produces 2 successes out of 30 attempts and on another system results in only 1 success out of 40, does not make the second system superior. The facility designer must examine the conditions and see how they apply to the particular installation at hand. Because this resistance to tamper estimate is so subjective, possibly its best use is to provide the system designer a list of weaknesses that an access control system exhibits so that steps can be taken to moderate or nullify these problems.

Conclusion

In the past, system designers have considered Type I and Type II errors in the selection of access control systems. It has been shown that quantitative estimates of these classical statistical errors are highly dependent on the likelihood of a particular penetration attempt on the facility. In the selection of an access control system, three performance estimates should be used -- user consistency, separability, and resistance to tamper, all of which can be obtained independently of the magnitude of the perceived threat. The facility designer is cautioned not to use these estimates as absolute fact, but examine how they were obtained before applying them with the threat in the performance equation to obtain the "best" balance between security, user acceptance, and the total cost of an access control system.

REFERENCES

1. A Papoulis, Probability, Random Variables and Stochastic Processes, McGraw-Hill, 1965

2. Entry-Control Systems Handbook, SAND77-1033, (Albuquerque: Sandia Laboratories, September 1977.)

3. Test Results Advanced Development Models of BISS Identity Verification Equipment, 4 vols, Report No. MTR 3442, MITRE Corp., September 1977.

4. C. N. LIU, et al, "Automatic Signature Verification: System Description and Field Test Results," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 1, January 1979.

5. Guidelines on Evaluation of Techniques for Automated Personal Identification, FIPS Pub. 48, April 1977.

6. G. R. Doddington, "Personal Identity Verification Using Voice," Proceedings: Electro-76, May 11-14, 1976, pp 22-24, 1-5.

7. G. H. Warfel, Identification Technologies, C. C. Thomas, 1979.