Human Factors in Software and Systems Engineering, Vol. 94, 2023, 133–144
https://doi.org/10.54941/ahfe1003779

A Formal Method for the Analysis of the Veteran's Ebenefits' Website

Giovanna Camacho, Matthew Bolton, Jingan Peng, Prashanth Wagle, and Lu Feng

University of Virginia, Charlottesville, VA 22903, USA

ABSTRACT

Currently, the Ebenefits/VA.gov website has login problems. This is a serious situation because it prevents veterans from accessing critical services: applying for medical disabilities; enrolling in health care services; accessing educational benefits; managing current VA benefits; acquiring home and auto loans, life insurance, burial services; and connecting veteran networks for other community resources. When a service member is unable to access this website, they are unable to help themselves. This adds inefficiencies to the overall VA system and could lead to poor veteran integration to civilian life. In this analysis, we used a technology called probabilistic model checking (a formal method for proving properties about stochastic systems) to identify the optimal process for veterans to login to Ebenefits while adhering to safety constraints for protecting veterans' sensitive information. To perform this analysis, a veteran in our research group documented multiple login attempts to gain realistic probabilities of the system transitioning between different interface states. Probabilistic model checking was used to quantify the probability of successfully getting from initial states to a successful login. In analyzing these results, the probability of a successful login for Ebenefits was found to be 0.25. Reviewing the data produced by the model checker revealed that a particular state called two factor authentication, utilized to verify the veteran's identity by a password and passcode sent to a technological device in their possession, was a problematic state. Our analyses also showed that one particular path, the Defense Self-Service Logon Path, was the most successful pathway at 0.98. This pathway begins with the user entering their password, verification of this password, followed up with a second authentication which the end user can skip or chose to add to their cellular device prior to allowing them to have a successful login. Based on this path, we found that use of a Common Access Card was most effective for enabling logins.

Keywords: Prism, Formal model specifications, Ebenefits

INTRODUCTION TO PROBLEM

Currently, the Ebenefits/VA.gov website: Home - VA/DoD eBenefits has login issues that prevent veterans from applying for medical disabilities, enrolling in health care services, accessing educational benefits, managing their current VA benefits, acquiring a home and auto loans, life insurance, burial services, as well as connecting to veteran networks for other helpful resources. When a service member is unable to access this website they are unable to self-help their processes which adds additional inefficiencies to the overall VA system.

VA service phone lines become inundated with long wait times, staff at local VA offices/hospitals are overwhelmed with frustrated emotional veterans needing help, and service members may get so frustrated they give up altogether. This can result in service members losing out on opportunities that could lead to unemployment issues or a rocky integration into civilian life.

Resolving this problem is imperative, as currently there are 19.4 million living veterans from a baseline census conducted on 30SEP20 who have access to this system (Veterans Affairs, 2020). These veterans range from 18 years all the way up to 112 years of age. All utilize this website and deserve to have easy access to their military benefits after their service to our country. This problem is interesting as it could offload a lot of inefficiencies in the health care system, help drive costs in processes down overall, and make both veterans and employees of veteran organizations happier with the end results. Fixing this website is long overdue. The entire website could use a better user interface design overall, but fixing the login issue alone would go a long way to allowing people the opportunity to access their benefits.

APPROACH

This project attempted to address this problem by using the power of formal methods. Formal methods are tools and techniques for modeling and proving properties about systems. Model checking is an approach to can do this automatically by rendering a system a searchable graph and then exhaustively searching it to find potential violations (Clarke et al., 1999). Probabilistic model checking, like that offered by the PRISM model checker (Kwiatkowska et al., 2011), also allows one to compute the probability of reaching certain states or prove stochastic properties. This research specifically follows in the tradition of efforts to use formal methods to evaluate human-machine interfaces (see a review in Bolton et al., 2013) while accounting for the stochasticity of branching paths (Bolton et al., 2021; Zheng et al. 2020a, 2020b).

The PRISM model checker was utilized to identify the optimal way for veterans to login to Ebenefits while also maintaining safety constraints for veterans' sensitive information. This specific objective was possible by reverse engineering. We analyzed the website's behaviors, deduced the underlying system model, and targeted the specifications that were causing current problematic issues when users attempt to log in.

Unfortunately, we did not have access to the underlying website's system model or specifications as this is a government-protected website. However, we do know that the website can be accessed by DS Logon, ID.me, and LOGIN.GOV. We also have a veteran within our group who was able to login various times to formulate the probability of success at the current state the website is in. This allowed the creation of the current system model from the end user's perspective.

A. *Analyze Website Behaviors With Formal Model Deduction*

The process allowed a better understanding of the human-machine interaction state and the ability to understand areas for improvement in enhancing its effectiveness and efficiency by targeting all elements depicted and reliant upon one another in Figure 1 that create the overall interplay of the system.

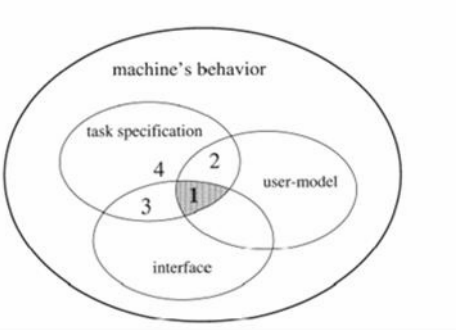


Figure 1: Degani and Heymann, 2002 formal model depiction of elements within a human and machine interaction.

Region 1 of this model with regards to our study delineates when correct interactions between the login user and the website interface have appropriate interactions. Region 2 is an example of an instance where the user may be completing the task requirements such as entering in their login information; however, the system is not completing its requirements for the task at hand appropriately login the user in. Region 3 of this model is an instance where the interface is correctly completing its requirements for the task such as sending the user a two-factor authentication (2FA); however, the user may not be completing this interaction appropriately (ex. completing the authentication at all or on time etc.). Region 4 is an area where both the user and the website are not completing the appropriate actions such as the user logging in incorrectly and then the website blocking the user immediately and sending a 2FA to (which is often automatically placed in a spam folder in their inbox).

B. *Target Specifications Causing Problematic Issues*

Further analyzing the problematic issues within the human-machine interface, we identified the following:

- Error states - user thinks they made a correct action, but the machine is driven to an incorrect state.
  - This occurs consistently within the two-factor authentication (2FA) portion of login process where the user thinks they are logging in, however, the website sends the user to various alternate states (ex. starting the process over, sending 2FA to spam mail, sending 2FA to the old phone number, etc.).
- Augmenting state – the user is told there is a mode, but the system has no such mode or cannot transition to the mode.
  - The user is told there are three different modes of logging in, but many times they are stuck and cannot transition out of intermediate screens.
- Restricting states – user is stuck in a mode.
  - This occurred frequently in Login.GOV where the user is stuck trying to retrieve the password sent to their phone and it does not send so they are stuck waiting leading to an unsuccessful login.

- Blocking states – the user is unaware of an event taking place
  - The user is unaware of whether their login is successful as the screen will reroute the login or kick them out completely of the website.
- Masked events – the event is not grouped into sets and not displayed individually.
  - The overall login method is grouped together and there are instances of confusion when various methods are using your phone, your email, or trying to get you to answer privacy questions. This leads to confusion.
- Unobserved events – these do not appear on the user's end of the model.
  - The biggest unobserved event is the success of verification utilizing all three modes of login.

Identifying these issues utilizing concepts introduced by Degani & Heymann (2002) helped to show that the overall website had severe login issues. Thus, the three various login methods needed to be further investigated to see what the best course of action was to help rectify the login processes for the end user.

MODELING

The best way to understand the various states within the login process to create a better PRISM model was to draw them out. The model in Figure 2 was created to help understand the flow of pathways occuring within the login system.

This models what is occurring when a user is trying to login to Ebenefits website. The user first begins at state 0 where they are to select an option that would be State 1 DS LOGON, State 3 ID.ME, or State 6. From the chosen state, the user then moves into State 2 Verify Credentials, State 4 Verify Credentials ID.ME, or State 6 Verify Credentials LOGIN.GOV. Once the user has their credentials verified, they all are sent to the two factor authentication (2FA) for a secure login. Once this step is completed, there can either be a login failure going to state 9 or there can be a successful login going to state 8. Once an individual has a login failure in state 9 they are usually

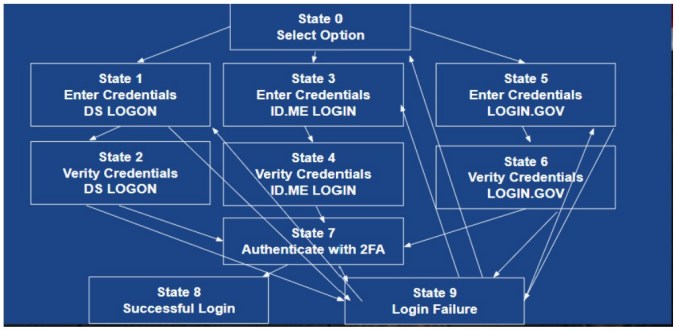


Figure 2: Login model depicting the states for logging into Ebenefits.

rerouted back to either the beginning at State 0 where they are to select a new option or they are routed back to enter credentials at whichever login path they chose. These 9 steps sound simple in nature on paper, but as just explained, there are various problematic issues within this model alone that need to be resolved to keep the specification errors explained earlier from allowing the following:

- Error States – State 7 to State 9 to ?
- Augmenting States – State 0 to State 1/State 3/State 5 to State 0
- Restricting States – Stuck in State 7
- Blocking States – Whichever state to State 0 or ?
- Masked events – Stuck in State 7
- Unobserved events – State 2/State4/State 6

There may be other states which we are not aware of as this is a government website; however, these are the basic states that we were certain would be useful in understanding probabilities down each pathway.

The Prism Model for this study was created by first identifying the previous states discussed. Then, the probability of success constants were created by a single user logging in consecutively ten times. From here the pathways were implemented in the formal model to understand what the probabilities were for success with three options. The pathway probability was later changed to only allow the functioning of each pathway specifically which will be further explained in our results and analysis section. However, the base code of our model along with the properties can be seen in the APPENDIX Figures 3 and 4, showing 16 states with 27 transitions.

RESULTS & ANALYSES

We used the PRISM model checker to evaluate the probability of successfully logging in to the website overall and via the different paths. The results of these analyses were as follows:

Probability of a successful login:	.25
Probability of an unsuccessful login:	.75
Probability of a successful DSL Path:	.98
Probability of an unsuccessful DSL Path:	.24
Probability of a successful IDM Path:	.12
Probability of an unsuccessful IDM Path:	.22
Probability of a successful Log Path:	.05
Probability of an unsuccessful LOG Path is	.28

In analyzing these results, the probability of a successful login for Ebenefits was found to be .25 with an unsuccessful login at .75. The DSL Path was the most successful pathway at .98. With results such as these, it would seem beneficial to consider eliminating the various pathways on the website and consider utilizing a single pathway that is working effectively.

Reconsidering each of the target specification issues, it is apparent that having multiple login methods has added confusion to the end user as they are not able to have enough transparency on what is occurring in each login

method. With older users who may not be as tech-savvy, these various login methods could add to the confusion. Having one stable login process could help simplify human-machine interface specifications and increase the current probability of login success.

Recognizing this as a problem to solve the situation, we reevaluated the model while treating each pathway as if each was the only option from the start.

The DSL pathway was found to have the following probabilities:

- Probability of successful DSL login: .28
- Probability of unsuccessful DSL login: .71

Looking at these probabilities it is apparent that the DSL login success had a slightly higher success rate (then a failure rate), but not substantial.

The IDM pathway was found to have the following probabilities:

- Probability of successful IDM login: .32
- Probability of unsuccessful IDM login: .67

Reviewing these probabilities, the IDM login had an even higher success rate, but also not substantial.

The LOG pathway was found to have the following probabilities:

- Probability of successful LOG login: .14
- Probability of unsuccessful LOG login: .86

The Log login probability had the lowest success rate.

DISCUSSION

Regardless of which pathway was chosen, this reducing login options did not fix all of the systems problems. However, the analysis of this intervention does show that utilizing Prism was helpful in evaluating alternatives. With this realization, multiple avenues could be considered to troubleshoot this website and fix other components to this issue and formally test them prior to completing an entire redesign of the website. Overall, the method can help to approach this problem in a more efficient manner which will help not only create faster changes, but also effective changes.

The probabilities created in this model were based on one individual's access to the Ebenefits website. However, if one user is having these issues, others are likely having them as well. Getting access to true login information was hard, so formatting the website's backbone of login was completed at a basic level. Furthermore, figuring out how to map the pathways within the code was something that had to be explained to our group as no one in our group was well-versed in formal methods, Prism.

It seems simplistic to suggest that the best way to begin addressing this website's login issues is to only allow for one login method. However, this does seem like a quick viable option to produce some change. The Ebenefits website already discusses some methods for addressing login

problems. These include clearing the cache, utilizing recommended browsers, utilizing activation codes, and recovering passwords. However, all of these methods are unreliable and place the work on the end user.

Mitigating all access issues should be completed through the engineering of the website as it is a service to a wide populace base. More than likely, there has been recent increased turnover in government IT. Reviewing the website and clearing messy code could improve the efficiency of this website. The user interface is extremely messy with various repetitive dead hyperlinks as is apparent with the 404 errors and redirecting occurring on the page. These errors can be monitored through Google Analytics, to help understand the problems that users are experiencing (Mandelbaum, 2021). When you can understand and monitor errors in real time, then you can act on them faster, helping to increase the functionality of the website. To maintain a cleaner website with high employee turnover, there should be a sitemap file that allows a clear road map of pages such as the login page and aid evaluations when updates are completed. Clearing out some of the subfolders for the login page could help simplify the URL string that users are trying to get to as well.

Tracking codes could help collect user data for those visiting the Ebenefits webpage. Acquiring more data on end users through human factors analyses could help to feed a formal method system that could help detect website problems. Problems addressed within the website can be placed in a formal methods model so that a simulation or automatically generated tests (Li & Bolton, 2019) could be run prior to updating the website to assure that the users are not subjected to more errors within the system.

Reviewing the specification states, it was also apparent that State 7 stuck out as a problem state with 2FA Authentication. The government must ensure that the appropriate users are accessing this website, as it does indeed have private veteran information that needs to maintain security. However, there needs to be an adequate balance of a compromise between security and functionality for users. Many veterans are not expert users and are consistently booted out of the system at the 2FA state.

With the goal of maintaining security but keeping the end users in mind, it would be helpful to evaluate alternative security options. There are three main ways in which security is acquired through something you know, have, and are (Stegnar, 2020). Something you "know," just means information such as a password, date of birth, maiden name etc. Something you "have," is like a mobile phone, or credit card. Examples of something you "are," can be voice recognition, biometrics, or fingerprints. The more difficult security is the something you are because it requires hardware to be present to support acquiring this information. When maintaining two factor authentication you are trying to acquire at least two of these three methods of security. This serves as a safety mechanism because if someone can get half of the information to break past the first barrier, they are still unable to retrieve the intended information. Furthermore, when individuals break past the first barrier the

end user is normally able to be notified so that they can fix the first layer of security that faltered.

The most important consideration in identifying security measures is to weigh the pros and cons within the three methods of security and figure out what methods work best for the system. Adding more layers of security can inhibit system processes much like what is being experienced in Ebenefits/Va.gov. The first problem with the something you "know" is that some individuals have bad passwords or security questions offered on the VA website aren't that hard for someone to look up and figure out. The next problem with the something you "have" is that the multi-factor authentication methods may go to an old device.

There may be issues if you do not have service within that area to receive the text verification, you broke your phone, or lost your device. The biggest problem with the something you "are" is that once the biometric data has been compromised, then it has been compromised for the rest of your life making it a noneffective method for security. In addition, individuals are less likely to want to give up their biometric data and much of the technology that common users utilize are not equipped with the expensive hardware to accommodate these security measures.

Alternative methods utilized to help individuals keep two step authentications with passwords secure, are by use of one-time passwords and time-based passwords. The one-time password utilizes a Hash-based Message Authentication Code (HMAC) of a secret key combined with a counter to generate a one-time code/password to enter the website. The time-based one-time password rounds the time to a reasonable time hack such as thirty seconds to replace the counter changing the password every thirty seconds to maintain a secure login. These methods are rather useful in helping to add the additional layer of security; however, this still does not alleviate the issues previously identified with regards to individuals needing to "have" the second known device to continue their logon process.

There are systems that attempt to make security easier for the end user by only requiring the two-step authentication when a new device is introduced. However, this method seems to backfire as the regular device will have saved passwords and information making the login process seamless and effortless for the user.

When the user is utilizing a newer device since they have not had to recall their password in some time, they may have forgotten their password or security questions. Many older veterans utilizing this system already may have issues with memory on trying to access these systems in the first place which adds to their frustration of added technology already.

There is still one very effective method that should be reconsidered for our veterans in reviewing the security and ease of use concerns. For years of veterans' lives on active duty they were required to carry around a Common Access Card (CAC). This CAC card requires a simple pin which allow verification of possession of the individual having this card and the password. This is a secure viable option as you require no extensive use of technology, knowledge, or memory of the end user. These users have already practiced

the habit of safeguarding this form of identification and are less likely to forget a pin than other information or hardware that would be required for a two-factor authentication. Should this CAC card be lost, then it would be a simple deactivation of the card which would protect the user's information. The biggest drawback to this CAC card would be requiring users to have the additional adapter required for their home computers to read this. However, this adapter is at a reasonable cost for ease of use and could most certainly be issued out with the CAC card with the money saved in health care efficiency with a better performing website.

FUTURE WORK

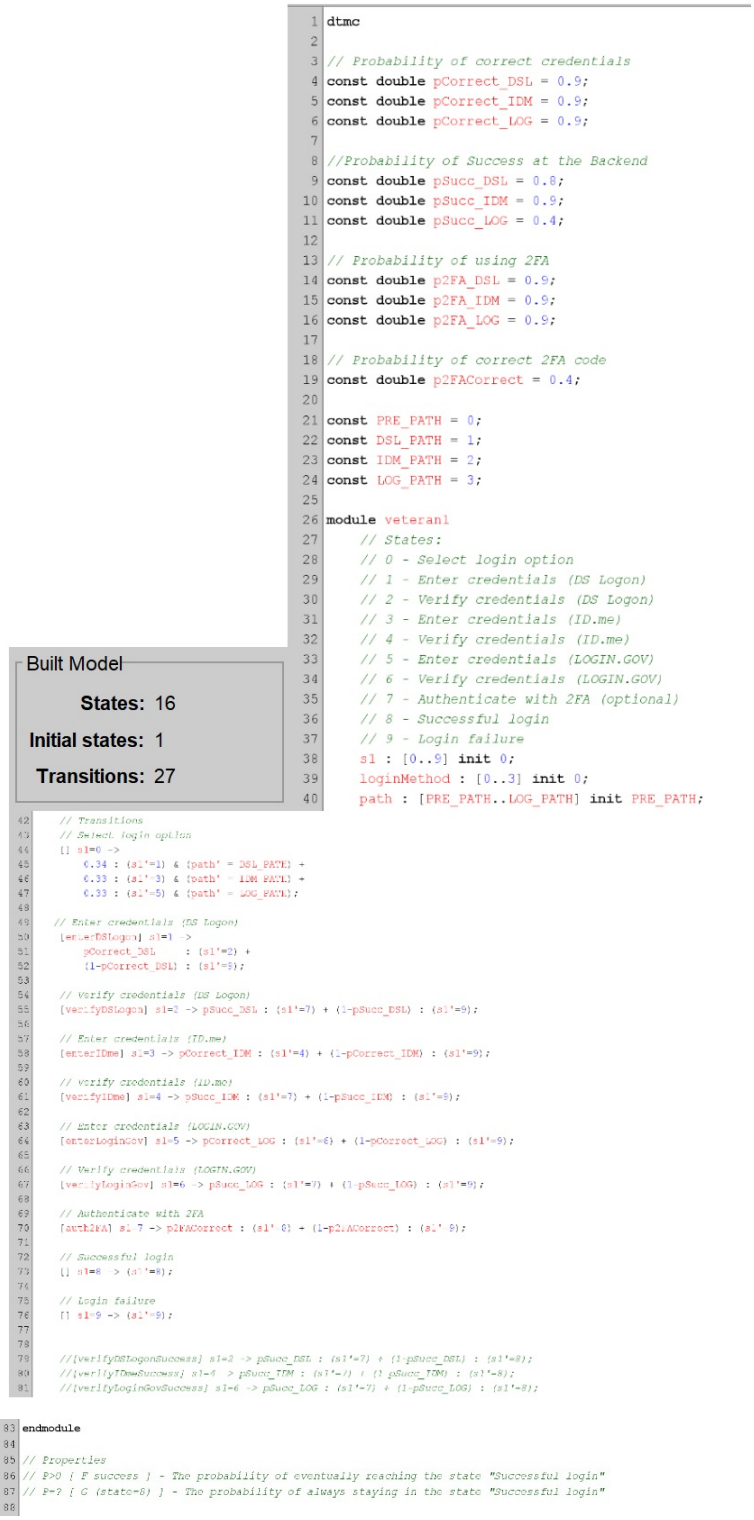
Further work must be completed utilizing formal methods to help implement changes to the Ebenefits website. In addition to gathering more user data, automated analyses could be used to compute accurate probabilities based on contextual factors (time of day, connection strength, geographical location). Automated analyses could be used to construct more complete formal models based on the website's source code. Understanding other problems outside of the direct user interface login could help to delineate other underlying issues. Furthermore, running a heuristic evaluation on end users can further help configure this website in an effective method for the end user.

Also, completing a test trial of individuals utilizing a CAC card to access their benefits is useful. Perhaps the biggest analyses that needs to be considered even prior to completing the work to revamp the Ebenefits/Va.gov would be to develop appropriate policies and procedures for the various benefits and services offered to veterans. Various veterans have described their experience with the VA very much like the following individual, "Logging on with the DOD (DS) identity is an exercise in futility. The overpaid bureaucrat that designed this should be tried and convicted for torture to military veterans" (Cheshire, 2021). Major processes overall within the system can be fixed beginning with this website. The motivation to fix this website should be our veterans. However, the definitive argument is very apparent that the cost of fixing this website is more than worth the time, as it will save the government much more money avoiding the inefficiencies of services on the backend.

Technology has advanced within the cyber physical system realm to enable us to give our veterans not only the health care that they need but the health care that they deserve.

APPENDIX

This model was created in conjunction with the following Prism Properties to find the probability of success and failure of each login pathway currently available in Ebenefits:



```
dtmc

// Probability of correct credentials
const double pCorrect_DSL = 0.9;
const double pCorrect_IDM = 0.9;
const double pCorrect_LOG = 0.9;

//Probability of Success at the Backend
const double pSucc_DSL = 0.8;
const double pSucc_IDM = 0.9;
const double pSucc_LOG = 0.4;

// Probability of using 2FA
const double p2FA_DSL = 0.9;
const double p2FA_IDM = 0.9;
const double p2FA_LOG = 0.9;

// Probability of correct 2FA code
const double p2FACorrect = 0.4;

const PRE_PATH = 0;
const DSL_PATH = 1;
const IDM_PATH = 2;
const LOG_PATH = 3;

module veteran1
    // States:
    // 0 - Select login option
    // 1 - Enter credentials (DS Logon)
    // 2 - Verify credentials (DS Logon)
    // 3 - Enter credentials (ID.me)
    // 4 - Verify credentials (ID.me)
    // 5 - Enter credentials (LOGIN.GOV)
    // 6 - Verify credentials (LOGIN.GOV)
    // 7 - Authenticate with 2FA (optional)
    // 8 - Successful login
    // 9 - Login failure
    s1 : [0..9] init 0;
    loginMethod : [0..3] init 0;
    path : [PRE_PATH..LOG_PATH] init PRE_PATH;

    // Transitions
    // Select login option
    [] s1=0 ->
        0.34 : (s1'=1) & (path' = DSL_PATH) +
        0.33 : (s1'=3) & (path' = IDM_PATH) +
        0.33 : (s1'=5) & (path' = LOG_PATH);

   // Enter credentials (DS Logon)
    [enterDSLogon] s1=1 ->
        pCorrect_DSL     : (s1'=2) +
        (1-pCorrect_DSL) : (s1'=9);

    // Verify credentials (DS Logon)
    [verifyDSLogon] s1=2 -> pSucc_DSL : (s1'=7) + (1-pSucc_DSL) : (s1'=9);

    // Enter credentials (ID.me)
    [enterIDme] s1=3 -> pCorrect_IDM : (s1'=4) + (1-pCorrect_IDM) : (s1'=9);

    // Verify credentials (ID.me)
    [verifyIDme] s1=4 -> pSucc_IDM : (s1'=7) + (1-pSucc_IDM) : (s1'=9);

    // Enter credentials (LOGIN.GOV)
    [enterLoginGov] s1=5 -> pCorrect_LOG : (s1'=6) + (1-pCorrect_LOG) : (s1'=9);

    // Verify credentials (LOGIN.GOV)
    [verifyLoginGov] s1=6 -> pSucc_LOG : (s1'=7) + (1-pSucc_LOG) : (s1'=9);

    // Authenticate with 2FA
    [auth2FA] s1=7 -> p2FACorrect : (s1'=8) + (1-p2FACorrect) : (s1'=9);

    // Successful login
    [] s1=8 -> (s1'=8);

    // Login failure
    [] s1=9 -> (s1'=9);


    //[verifyDSLogonSuccess] s1=2 -> pSucc_DSL : (s1'=7) + (1-pSucc_DSL) : (s1'=8);
    //[verifyIDmeSuccess] s1=4 -> pSucc_IDM : (s1'=7) + (1-pSucc_IDM) : (s1'=8);
    //[verifyLoginGovSuccess] s1=6 -> pSucc_LOG : (s1'=7) + (1-pSucc_LOG) : (s1'=8);

endmodule

// Properties
// P>0 [ F success ] - The probability of eventually reaching the state "Successful login"
// P=? [ G (state=8) ] - The probability of always staying in the state "Successful login"
```

Built Model

States: 16

Initial states: 1

Transitions: 27

Figure 3: Prism model created from the states identified along with their respective pathways.

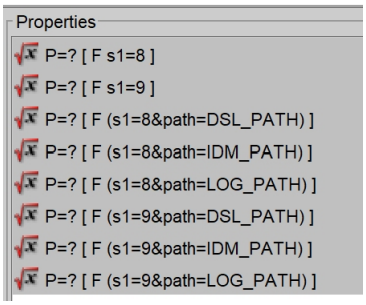


Figure 4: Prism properties created to attain probabilities of login success and failure down various login pathways.

ACKNOWLEDGMENT

This paper was created for a Formal Methods, Safety and Security class project at the University of Virginia.

REFERENCES

Bolton, ML, Bass, EJ, Siminiceanu, RI. Using formal verification to evaluate human-automation interaction in safety critical systems, a review. IEEE Transactions on Systems, Man and Cybernetics: Systems. 2013, 43(3), 488–503. doi: 10.1109/TSMCA.2012.2210406.

Bolton, ML, Zheng, X, Kang, E. A formal method for including the probability of erroneous human task behavior in system analysis. Reliability Engineering & System Safety, 213, 13 pages. doi: 10.1016/j.ress.2021.107764

Clarke, EM, Grumberg, O, Peled, DA. Model checking.1999, MIT Press.

Degani A, Heymann M. Formal verification of human-automation interaction. Hum Factors. 2002 Spring;44(1): 28–43. doi: 10.1518/0018720024494838. PMID: 12118871.

Heimdahl, Mats P. E., Formal Methods for Developing High Assurance Computer Systems: Working Group Report. Defense Technical Information Center, http://discover.dtic.mil/.

Li, M, Bolton, ML. Task-based automated test case generation for human-machine interaction. Proceedings of the International Annual Meeting of the Human Factors and Ergonomics Society. 2019, 807–811, Sage: Thousand Oaks. doi: 10.1177/1071181319631157.

Mandelbaum, Aaron. "Top 12 Most Common Website Problems and How to Fix Them." Paradox Marketing, 20 Oct. 2021, https://paradoxmarketing.io/capabilities/search-engine- optimization/insights/top-12-most-common-website-problems-and-how-to-fix-them/.

Kwiatkowska, M, Norman, G, Parker, D. PRISM 4.0: Verification of Probabilistic Real-time Systems. In Proc. 23rd International Conference on Computer Aided Verification (CAV'11), 2011, 585–591.

Planning, Office of Policy and. "Veterans Affairs." Go to VA.gov, 24 Nov. 2010, https://www.va.gov/vetdata/additional_sources_of_information_about_veterans.asp.

Stegner, Ben. "The Pros and Cons of Two-Factor Authentication Types and Methods." MUO, 15 Apr. 2020, https://www.makeuseof.com/tag/pros-cons-2fa-types-methods/?newsletter_popup=1.

"Va.gov Site Becomes Central Login for Accessing Benefits." VA News, 29 Apr. 2021, https://news.va.gov/88099/va-gov-site-becomes-central-login-accessing-benefits/.

Vetpop2020: A Brief Description - Veterans Affairs. https://www.va.gov/vetdata/docs/Demographics/New_Vetpop_Model/VetPop2020_A_Br ief_Description.pdf.

Zheng, X, Bolton, ML, Daly, C. Extended SAFPHR (Systems Analysis for Formal Pharmaceutical Human Reliability): Two Approaches Based on Extended CREAM and a Comparative Analysis. Safety Science. 2020a, 132, 18 pages. doi: 10.1016/j.ssci.2020.104944.

Zheng, X, Bolton, ML, Daly, C, Biltekoff, E. The development of a next-generation human reliability analysis: Systems analysis for formal pharmaceutical human reliability (SAFPHR). 2020b, Reliability Engineering & System Safety, 202, 15 pages. doi: 10.1016/j.ress.2020.106927.