Behavioural Biometrics and Cognitive Security Authentication Comparison Study

Karan Khare[1], [2]Surbhi Rautji and [3]Deepak Gaur

[1]Gradestack Pvt. Ltd, TLABS, Times Tower, Noida (UP), India
[2]Oracle Technologies, Oracle 3C Building, Sector 127, Noida (UP), India
[3]Faculty, Department of Computer Science and Engineering, AMITY University, Noida, India.

ABSTRACT

Behavioural biometrics is a scientific study with the primary purpose of identifying the authenticity of a user based on the way they interact with an authentication mechanism. While Association based password authentication is a cognitive model of authentication system.

The work done shows the implementation of Keyboard Latency technique for Authentication, implementation of Association Based Password authentication and comparison among two. There are several forms of behavioural biometrics such as voice analysis, signature verification, and keystroke dynamics. In this study, evidence is presented indicating that keystroke dynamics is a viable method not only for user verification, but also for identification as well. The work presented in this model borrows ideas from the bioinformatics literature such as position specific scoring matrices (motifs) and multiple sequence alignments to provide a novel approach to user verification and identification within the context of a keystroke dynamics based user authentication system. Similarly Cognitive approach can be defined in many ways of which one is association based Technique for authentication.

KEYWORDS

Keyboard Dynamics, Behavioural Biometrics, Association based passwords, keyboard latency, cognitive authentication, Behavioural authentication.

1. INTRODUCTION

With the increasing number of E-commerce based organizations adopting a stronger consumer-orientated philosophy, web-based services (E-commerce) must become more user-centric. As billions of dollars worth of business transactions occur on a daily basis, E-commerce based enterprises must ensure that users of their systems are satisfied with the security features in place. As a starting point, users must have confidence that their personal details are secure. Access to the user's personal details is usually restricted through the use of a login ID/password protection scheme. If this scheme is breached, then a user's details are generally open for inspection and possible misuse. Hardware (physiological) based systems are not yet feasible over the Internet because of cost factors and in addition, the question as to their ability to reduce intruder detection has not yet been answered equivocally. One thing is for certain, providing every household with a retinal scanner and instructions on its usage has not yet reached mainstream society. The extent to which hardware based security enhancement systems are able to reduce the imposter acceptance rate is still study dependent and the results indicate that the false acceptance ratio (FAR) is still on

the order of 5%, beyond the acceptable risk level of many organizations (and individuals) considering the costs in terms of hardware and training time. We propose an inexpensive (virtually free) software based enhancement to class C (login ID/password) security measures that provides a cross-over error rate with respect to false acceptance/false rejection ratios that is very competitive with hardware based systems both in terms of accuracy and monetary outlay.

The system is based on what has now become known as "keystroke dynamics" with the addition Keyboard Latency Approach. We also consider in this study the affect of typing speed and the use of a rhythm when a user enters their login details. Keystroke dynamics was first introduced in the early 1980s as a method for identifying the individuality of a given sequence of characters entered through a traditional computer keyboard. Researchers focused on the keystroke pattern, in terms of keyboard duration and keyboard latency. Evidence from preliminary studies indicated that when two individuals entered the same login details, their typing patterns would be sufficiently unique as to provide a characteristic signature that could be used to differentiate one from another. If one of the signatures could be definitively associated with a proper user, then any differences in typing patterns associated with that particular login ID/password must be the result of a fraudulent attempt to use those details. Thus, the notion of a software based biometric security enhancement system was born.

The work also represents the use of Association Based Passwords. It rests on the human cognitive ability of association-based memorization to make the authentication more user-friendly, comparing with traditional textual password. Based on the principle of zero-knowledge proof protocol, The Model further improve the primary design to overcome the shoulder-surfing attack issue without adding any extra complexity into the authentication procedure. System performance analysis and comparisons are presented to support my proposals.

2. THE TWO MODELS: BASIC IDEA

Textual passwords has been a decade now since they have been implemented almost everywhere but the need arose to use some latest technology since the text passwords started suffering hackers attack thus causing a heavy loss of information to the organization and the individuals.
Thus there was a need to technology which limits the identity of the person to himself only and getting his id by the hackers and the intruders is almost impossible. There came than the use of BIOMETRICS and COGNITION BASED SYSTEMS.

Biometrics can be further classified as Physical Biometrics and Behavioural Biometrics. Physical Biometrics deals with various physical parts of the human for identification like, Retina Scan, Fingerprint etc. While the later i.e. Behavioural Biometrics deals with human behavioural approach like Keyboard Dynamics, Mouse Dynamics, Signature Dynamics etc. Cognitive model of security and authentication on the other hand use the approach towards human brain memorization and association powers.

Here out of the two classes' one approach each has been considered and is compared to get better accuracy and authentication mechanism. The two models being Keyboard Dynamics based on Keyboard Latency Test and Association Based Password System. The two have been considered for implementation since they are new in the area and does have many pattern to suffer hackers attack since the responses for authentication is either related to human brain or to his behavior and stealing one brain capability and behavior is impossible. Yes but copying it by means of some hardware is definitely possible. More over these two approaches does not require any

special Hardware for the purpose of Authentication thus making it a more reliable, efficient and negligible investment systems and operating them is also a fun.

3. DESIGN AND IMPLEMENTATION

The primary goal of the software system is to implement the following modules advanced security module that requires no extra hardware, investment and is easy to implement on a standalone machine as well as web based application. Thus for the same reason the software system consist of two basis approaches as one system.

The two approaches implemented here are Keyboard Dynamics based on Keyboard Latency and Association Based Passwords. Both of the techniques use a common feature of maintaining a profile at backend.

Before explaining actual design phase the three hypotheses was taken into consideration, these three hypotheses were:

I. Null Hypothesis

The practice of science involves formulating and testing *hypotheses*, assertions that are capable of being proven false using a test of observed data. The **null hypothesis** typically corresponds to a general or default position. For example, the null hypothesis might be that there is no relationship between two measured phenomena or that a potential treatment has no effect.

Here the approach to Keyboard Dynamics based on Keyboard Latency is been considered under **null hypothesis** since it also can be considered under the tendency to be proven wrong for authentication.

II. Can Latency Values be used for authentication?

For considering the Keyboard Latency as approach to authentication it needed to be proven first that the Latency values can actually be used for the authentication of the user and does not suffer any shoulder attack.

III. Is **Keyboard Latency** based authentication a better approach to Characteristics based password i.e. here **Association Based Passwords**?

Characteristics or the Association Based Passwords can be a challenging model to Keyboard Latency based test, a strong comparison between Behavioural Biometrics Model and a Cognition Model of authentication system.

3.1 Approach to Design

3.1.1 The Keystroke Dynamics

The basic idea over the implementation of the system can be made out from the figure as illustrated in Fig1, where client connects to the system where the Logic is stored for Authentication and various other databases are used to implement that logic and provide data when required. The implementation that will be shown here is basically for testing the correctness

of the system and thus is implemented on a Standalone machine but this basically should be used in a network environment.

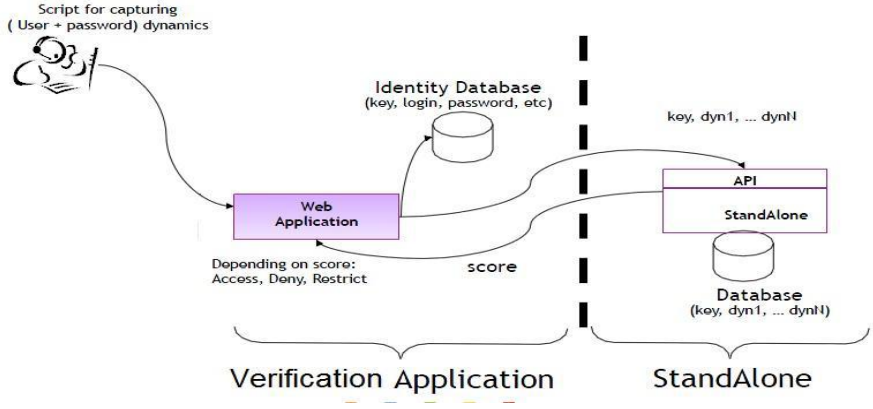

Fig 1: Architecture for Behavioural Biometrics Keyboard Dynamics

Keystroke dynamics is a particular instance of a behavioural biometrics that captures the typing style of a user. The dynamics of a user's interaction with a keyboard input device yields quantitative information with respect to dwell time (how long a key is pressed) and time-of-flight (the time taken to enter successive keys). By collecting the dynamic aspects acquired even during the login process, one can develop a model that captures potentially unique characteristics that can be used for the identification of an individual. To facilitate the development of the model of how the user enters their details, an enrollment phase is required, when the user is asked to enter his/her login id/password until a steady value is obtained (usually limited to 5-10 trials - but this is implementation dependent). Once this data has been collected, a reference 'signature' is generated for this user. The reference signature is then used to authenticate the user account on subsequent login attempts. The user with that particular login id/password combination has their keystroke dynamics extracted and then compared with the stored reference signature. If they are within a prescribed tolerance limit – the user is authenticated. If not – then the system can decide whether to lock up the workstation - or take some other suitable action.

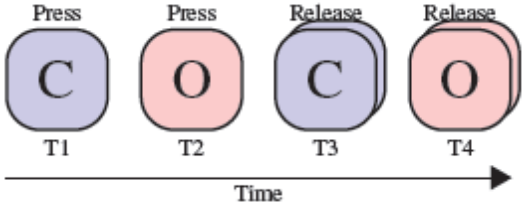

Fig 2: Dwell Time calculation

When devising such a biometric solution - there is always a tradeoff between being overly stringent - rejecting every attempt to login in and being overly lenient - allowing imposters to access the computer. The former is usually reported as a measure of false rejection - a type I error and the later a false acceptance or type II error. Another measure - called the cross over error rate (CER) - sometimes referred to as the equal error rate (EER) is also reported which provides a quantitative measure of how sensitive the biometric is at balancing ease of use for the authentic user while at the same time reducing the imposter access rate. All extant biometric systems yield a trade-off between these two measures - those that reject imposters effectively (low FAR) are

usually accompanied by a high FRR (rejection of authentic users) and vice versa. The next section presents a brief description of the methodology employed, followed by a results and conclusion section.

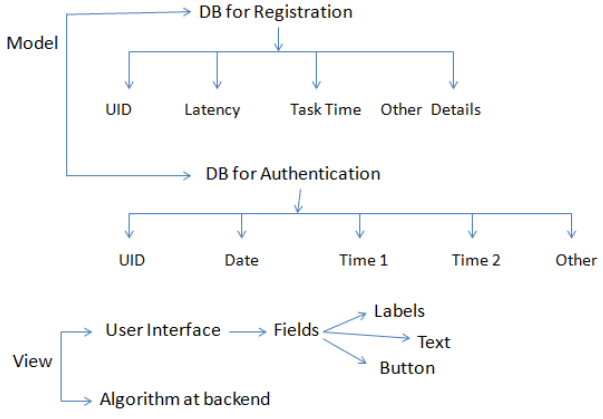

Fig 3: Design approach to keyboard latency system

3.1.2 The Association Based Password

This password system is cognition model in this the user first registers himself for authentication process to be followed. During registration the user registers himself with USERNAME and PASSWORD entries. These password entries are basically a user's association, association that first comes to the user mind. Example 1: If a field put in front of the user is **sky** than whatever comes into the mind of the user on relating it, he puts up that in the Password field let's say **stars.** The user in the registration process enters few password entries (here considered 6 fields). After this is entered a USER profile is maintained in the database, this profile is the original profile and cannot be changed without verification of the USER.

Now after the registration is done the next window that is entertained in this prospect is the password verification window. In this password verification window the USER fist enters his Username and random entries of the password fields Generated out of the total password entries (here considered 2 out of 6). After the user enters the entries in the window than on submitting the entries for authentication a temporary profile is created. This temporary profile is then compared with the original profile. After a successful verification the control access is given to the user.

This association based technique makes use of human cognition and protects the password from various attacks like shoulder surfing attack, pattern attack etc. The Results for this approach is discussed in the **RESULTS** section.

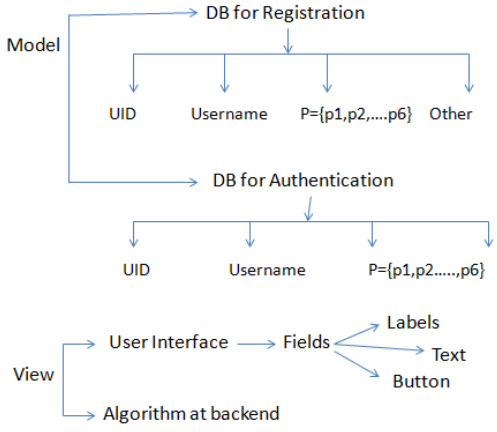

Fig 4: Design Approach to Association based passwords

4. SIMULATION

4.1 Keyboard Latency approach

The first and foremost step for the user is to create a registration profile, since verification only produces result when there is original profile is created for the USER. If there is no original profile created for the USER than the Latency or Association Test will not produce any result [which will produce a system hang after for few minutes of 5 unsuccessful attempts].

When registering, user can input combination of passwords i.e. numeric, alphanumeric and alphabetical. Here the numeric, Alphabetic and Alphanumeric specifies entries restricted to numeric form, Alphabetic form and combinations of Alphabets and Numeric characters. This is done so that the Latency can be computed in various aspects since USER accessibility depends on the keys pattern and it is taken under consideration. The user will probably take more time in typing numeric password on laptop keyboard where numeric keys are in a single row, while it will take least time in Alphabetic sequence.

After the USER register with the password which is limited to 6 characters per field which when submitted stores up in the same table where USERNAME was getting stored up. Now If the USER Selects the Practice option than for him it is necessary that he should be a registered USER i.e. his USERNAME and PASSWORD details should exist already than only his latency score can be stored up or else there will be no field as such to store up the Latency Score for the USER. And also is necessary and among the most important step. Since this is the only step which is responsible for calculating the DWELL TIME. The practice Session for a particular USER is done 5 times and the average of 5 Practice sessions is calculated and since it is not necessary the latency for the first time is same as that of the last time and that actually depends on the expertise of the USER to interact with the keyboard. Thus this step is most important of all the steps in Keyboard Latency Technique.

The record is made for respective entries i.e. numeric, alphanumeric and alphabetical occurrences and passwords and the difference of time between key press and key release, time between two consecutive keys and total time latency. And finally the profile gets created for the user.

At the time of actual authentication the various fields, i.e. username and password are taken as input by the user and respective Latency is calculated and are recorded for Numeric Verification, Alphabetic Verification, and Alphanumeric Verification respectively and this is the temporary profile created for the USER. This temporary profile created is when compared with the original profile there is a condideration inclusion of a **PARAMETERIC** value **P.** This is the parametric value which differs for individual key latency of both Numeric Alphabetical and Alphanumeric entries. This value of **P** is calculated with the survey done on a group of 100 people.

Let's say the Latency value is Lat1 than it will be considered as **Lat1 (-+) P** and than this calculated with the original profile value **LAT.** Which than incase produces the result and provides the Authentication.

4.2 Association Based Password Approach

Now the other module represents the **Association Based Passwords as application of cognitive model**. In this model first and foremost the USER registers himself for the Association Based Password Testing. The USER enters the USERNAME and the Set of fields for specifying the Password with Association. This can be seen in

After the Registration is Done User can select the Option for Verification, here the USER enters the verification information i.e. the USERNAME, and 2 password fields randomly generated these should have the values that were associated earlier should be same this time also these two fields generated can be any of the fields out of the 6(considered in this work) and should have the same responses as specified earlier.

Example Illustration:

Alice wants to authenticate herself with the Association based password so the steps she follows:

1. Alice registers herself in the Registration window.
2. Alice enters her USERNAME and PASSWORDS.
3. USERNAME is a single field and Alice enters ALIS546.
4. Now Alice enter PASSWORD set p= {p1, p2, p3, p4, p5, p6} i.e. 6 association based passwords for herself.
5. The Password Fields are p= {sky, world, college, politics, bike, coffee}.
6. Alice enter Password answers Pa= {stars, earth, friendship, bad, ride, costa}.
7. Next Alice clicks submit button to create the original profile.
8. Now Alice interacts with verification window.
9. Files generated before her are: F= {USERNAME, COLLEGE, BIKE}.
10. Alice enter FA= { ALIS546, friendship, ride}
11. The authentication is provided to ALICE.

5. Results

The result analysis was done for Keyboard Dynamics and Association Based Passwords for 100 people.

The keystroke sequences are corresponding to user names and passwords with fixed length of 6 characters. The data base is also containing impostor's attacks for each user. Each user has

provided between 20 and 110 logins sequences and some people have been asked to try to reproduce some sequences between 20 and 100 times. The different methods proposed to adapt the parameters (the security threshold and the fusion weights) for each user have been evaluated by using the leave one out method. It has been estimated the parameters of one user with a tool trained on all the other users. Implementation of real life applications should also integrated our private database. This database will be considered as a training set and is supposed to be representative of the different classes of users. Results obtained are presented in table.

Table shows important improvements compared to the use of global parameters. Performances are improved for all the classes. The obtained error rates are very good for a keystroke dynamics method. However, these error rates hide the fact that the error is computed on all profiles of a class. It tends to minimize the influence of low performance users, who has catastrophic results. We have identified three of this type of users in our base (EER>30%). If we compute the average of the EER computed on each user we obtain 4.5%, corresponding to a fair performance. This value points another problem of our method: probably, because of the few numbers of problematic users, we are unable to achieve our second objective which was identifying them before the authentication with our clustering methods.

	FAR% Global Parameter	FRR% Global Parameter	FAR %	FRR%
class 1	0	0.1	0	0
class 2	1.8	5.8	2.8	3.3
class 3	0	1.9	0	1.9
Total	1.8	5.3	1.7	2.1

Table1: Results on the Basis of Classification for Keyboard Dynamics

Similarly the User study was conducted for Association Based Passwords and the study was again carried out for 100 USERS and the Analysis met on an average 4 correct responses out of 5 producing accuracy rate of 80% thus stating that the efficiency of Association Based password is good but a bit complex for USERS and a sample of survey states that the USER does not associate all the time thus producing a incorrect response. This part of the Survey was conducted for each USER 5 times after and increasing period of 3 days, 5 days, 7 days, 10 days and 15 days. Normally it was found out that It faces association problem with the passage of time and It becomes a bit complex for the USER to keep remembering the answers or associations to the response and generally the association based responses comes out based on the mood of the USER thus facing a problem while responding after a long time.

6. CONCLUSION

In our work we showed that the security models can be much more effective when combined with behavioral or cognitive abilities of the user. For this context, our method outperforms all other tested methods from the state of the art. For keyboard latency it has been observed that the

individual threshold has improved the overall performance of the system while in the case of the association based passwords it has been observed that the performance is no of comparisons dependent i.e. more the number of association profiles maintained higher will be the performance overhead.

In our study it was also found out that the Keyboard Dynamics gave the better response in comparison to Association based password according to response of 20 persons. Since the study was restricted to 20 users, so a study over a wider area and among more users may give different results which are exactly our future plan of study. Also the consideration of parametric value here made according to study over an above specified amount of users so there is still a chance of a minimal change in the precision.

The length in the password fields of Keyboard Dynamics and Association based passwords is kept fixed when these tested with variable length may produce different results. Probably it may be possible that if the number of fields in the Association Based Passwords is increased and more number of Association tests is provided than this may produce some different results.

7. FUTURE SCOPE

The Software and the implementation has a very high future scope of development since the application is developed under limited circumstances and need to be tested in a wide range of circumstance which entirely is dependent on the no of USER participants.

The application and the field has the tendency to evolve further and this can be achieved only when the response is collected among a variety of users classified among different categories like some mentally weak, persons with low typing speed, the latency may be dependent over age of the user and may vary with his age and change of keyboard configuration.

There remain a number of Future implementations for this area and they can be classified as:
Apart from Keyboard Dynamics Behavioural Biometrics also has the scope for Mouse dynamics, Signature Dynamics, Speech Dynamics etc. Thus these implementations can also be done in the project. There may be a change in the value of Parametric Variable **P** since the study right now conducted was with just 100 people while this result can vary highly when conducted with 300 people. And this might produce the higher accuracy rate for the System. Right now the Keyboard Dynamics model is not Keyboard dependent and with some improvements and algorithms this can be made Keyboard independent and flexible. Many things like the latency when user stays still etc have been considered here in this model in ideal basis but need to be implemented on practical analysis basis. The application is been completed in the standalone environment and not been done in the network environment so there remains a scope for this in network field as well.

REFERENCES

[1] Yan, J., Blackwell, A.F., Anderson, R. & Grant, A., 2004, Password memorability and security: Empirical results, IEEE Security and Privacy 2(5), 25-31.

[2] Alen Peacock, Xian Ke, Matthew Wilkerson. "Typing Patterns: A Key to User Identification," IEEE Security and Privacy, vol. 02, no. 5, pp. 40-47, September-October 2004.

[3] Joyce, R. and Gupta, G., 1990. Identity authorization based on keystroke latencies. Communications of the ACM. Vol. 33(2),pp 168-176.

[4] Monrose, F. et al, 2001. Password Hardening based on Keystroke Dynamics. International Journal of Information Security.

[5] Zhi Li, Qibin Sun, Yong Lian and D. D. Giusto, "An Association-Based Graphical Password Design Resistant To Shoulder-Surfing Attacks", IEEE International Conference on Multimedia and Expo (2005) , 245-248.
[6] S. Cho and S. Hwang, "Artificial rhythms and cues for keystroke dynamics based authentication," in IAPR International Conference on Biometrics, vol. 5, 2006, pp. 626–632.
[7] E. Yu and S. Cho, "Keystroke dynamics identity verification—its problems and practical solutions" Computer & Security, vol. 23, no. 5, pp 428-440, 2004
[8] B. Schneier, Applied Cryptography, New York: Wiley 1996.
[9] V. Roth, K. Richter, R. Freidinger, "A PIN-Entry Method Resilient Against Shoulder Surfing", 1th ACM Conference on Computer and Communications Security (CCS'04), Washington DC, USA, Oct, 2004.
[10] S. Hocquet, J.-Y. Ramel, and H. Cardot, "User classification for keystroke dynamics authentication," in The Sixth International Conference on Biometrics (ICB2007), 2007, pp. 531–539.
[11] R. Giot, M. El-Abed, and R. Christophe, "Greyc keystroke: a benchmark for keystroke dynamics biometric systems," in IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS 2009), 2009, to be published.
[12] D. Hosseinzadeh and S. Krishnan, "Gaussian mixture modeling of keystroke patterns for biometric pplications," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol. 38, no. 6, pp. 816–826, 2008.
[13] V. V. Phoaha, S. Phoha, A. Ray, S. S. Joshi, and S. K. Vuyyuru, "Hidden markov model (hmm)-based user authentication using ystroke dynamics," patent, fev 2009.
[14] G. Ginesu, D. Giusto, T. Onali, "Image Based Authentication (IBA): A Review", N3461, ISO/IEC JTC 1/SC 29/WG1, Nov, 2004.
[15] D. Song, D. Brumley, H. Yin, J. Baballero, I. Jager, M.G. Kang, Z. Liang, J. Newsome, P. Poosankam, and P. Saxena. BitBlaze: A new approach to computer Security via binary analysis. In proceedings of International Conference of Information Systems Security, Hyderabad, India, 2008.
[16] Edward J. Schwartz, Thanassis Avgerinos, and David Brumley. Q: Exploit Hardening Made Easy. In Proceedings of the USENIX Security Symposium, 2011.
[17] Jager, T.Avgerinos, E. Schwartz, and D. Brumley. BAP: A Binary Analysis Perform. In proceedings of the Conference on Computer Aided Verification, 2011.

Authors

Karan Khare is Web Developer and a security specialist in Gradestack Pvt. Ltd. He has completed his Bachelor of Technology in Computer Science and Engineering from Amity School of Engineering studies. Karan holds research interest in Artificial intelligence, Genetic Algorithms, Image Processing, Cognitive studies and Computer Network Security.

Surbhi Rautji is a Software Engineer in Oracle Technologies. She has completed her Bachelor of Technology in Computer Science and Engineering from Amity University, Noida, India. Her fields of interest include Image Processing, Image security, Artificial Intelligence, Genetic Algorithms and Computer Network Security.

Deepak Gaur had received Master of Engineering in Computer Science & Engineeringfrom Punjab Engineering college, University of Technology, Chandigarh. He hascompleted his B.Tech in Computer Science & Engineering from Himachal Pradesh University, Shimla(H.P). Presently he is working as Assistant Professor in CSE Department, Amity University, Noida, Uttar Pradesh, India. Mr. Deepak Gaur, research Area is Image Processing, Image Compression, Security Systems, Image Security and pattern reorganization.