American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-11, Issue-03, pp-23-32
www.ajer.org
Research Paper Open Access

A Data Security Model for Preventing Mass Data Breaches

Jianqing Wu, Ph.D., J.D., Ping Zha, M.D.
Correspondent Author: Jianqing Wu, Ph.D., J.D.
***Bios of the author**: Jianqing Wu, Ph.D., J.D. has developed internet data storage security technologies for nearly two decades with four granted U.S. Patents in data storage security. He also does pioneer medical researches to find cures for chronic diseases.*

Abstract
This article deeply explores the root cause of mass data breaches and a new approach to preventing future data breaches. The prevalent mass data breaches can be attributed to the existing data security model adopted to early network computers. By following this model, consumer data records are stored in the same way. If the security measures of a network server are cracked, all consumer data records in different user accounts can be stolen in a similar way. The potential rewards from a successful hacking are much larger than the cost needed to crack the security measures. In such a system, whatever data security measures can be defeated by insider help which can be secured by sharing only a small portion of rewards. Due to unavoidable risks of insider assistance, the highest consumer data security cannot be achieved by merely improving technologies, but can be achieved by altering data security ecosystem. I will show that encrypting consumer data by the consumer's own keys in a semi-trust business relationship model will dramatically reduce hacking incentive and thus prevent future data breaches. I will discuss the conditions for using such encryption technologies and great benefits for addressing data security for the U.S. residents, whose identity data have been stolen.

Key words: *cyber security, mass data breach, data encryption, secured business transactions, identity theft*

Date of Submission: 02-03-2022 Date of acceptance: 17-03-2022

I. INTRODUCTION

Computer security was developed based on early data security model that was focused on the computer and the data owned by the owner of the computer [1]. In the early time, user data may be stored on a network computer incidentally. If user data are stored on a network computer, they most probably did not have ascertainable value. Naturally, data security model and security measures were developed by focusing on only the computer with almost no attention to data ownership. Early studies found the importance of trust in buyer-and-sell transactions unrelated to e-commerce [2-4]. When E-commerce appeared, it is recognized that successful business relationship between a seller and a buyer depends on mutual trust. For a vendor to succeed, the vendor must win trust from consumers in the competitive world [5-12]. In order to maintain continuous business relationship, vendors need consumer information. For convenience, it is desirable to have consumer data stored on the server so that the vendor can use them whenever a need arises. The vendor takes personal information, address, banking information, credit card, etc. and keeps the data in the same way. When the vendor becomes larger and larger, the magnitude of risk of leaking consumer data also increases. This is why we see endless mass data breaches. Wikipedia documented about 349 mass data breaches [13-14]. Most of small data breaches may not get attention and thus escape from scrutiny. The database of Privacy Right lists more than 9,000 data breaches [15]. One worst data breach with the MEGA cloud service exposed 772,904,991 emails and 21,222,975 passwords [16]. Over the decades, it is hoped that mass data breaches can be stopped by improving security measures. However, data security measures work only marginally because they can be cracked for the same reason. In this study, I will show a workable data security measure must depart from the existing data security model and change the data security ecosystem.

II. FLAWS IN EXISTING DATA SECURITY MODEL

A. Prevalent Consumer Data Breaches

Internet history reflects failure in protecting consumer data. The Wikipedia shows a list of 349 well known mass data breaches for those involving the theft or compromise of 30,000 or more records and those involving large corporations [13]. It was estimated that the total damage caused by data breaches is between $375 and $575 billion annually in the U.S. The global annual cost forecast is about $2.1 trillion. According to the non-profit consumer organization Privacy Rights Clearinghouse, a total of 227,052,199 individual records containing sensitive personal information were involved in security breaches in the United States between January 2005 and May 2008, excluding incidents where sensitive data was apparently not actually exposed [15]. In 2019, a collection of 2.7 billion identity records, consisting of 774 million unique email addresses and 21 million unique passwords, was posted on the web for sale [16]. The total number of consumer records exposed in 2021 is at least 375 million which is based on the best estimates (excluding those without providing the number of consumer records). The accumulative amount of consumer data loss is extremely large.

Despite a large number of patent grants in the U.S. alone, there is no sign that existing technologies can ever stop mass data breaches. Data breaches may happen to any companies such as Facebook, Yahoo, Sony, Google, and T-mobile, and many of them experienced data breaches multiple times. A large number of mass data breaches happened to banks, credit card companies, credit reporting agencies, health care providers, and schools and universities. For 2021, 13 mass data breaches were reported. Based on the number and pattern of data breaches and the technical and financial strengths of the involved companies, I can reasonably conclude that an effective solution may be found only by seeking a new data security model.

The impacts of breached consumer data may be enhanced by unregulated access to data broker market [17]. Data broker is now more than $200 billion industry [18]. A data broker collects consumers data from federal government (the U.S. Census Bureau), state agencies, retailers, other data brokers, and internet companies [19-20]. One of the companies the FTC studied, Acxiom, boasted that “[i]ts databases contain information about 700 million consumers worldwide with over 3,000 data segments for nearly every U.S. consumer.” [19]. If I combine the consumer data from mass data breaches and consumer data traded in the data broker industry, I must find that few U.S. residents can keep their identity data.

B. Two Incentives for Data Breaches Results in Active Hacking Trade

Consumers data may be sold for two different purposes: some data are sold as market leads and others sold for defrauding consumers, governments, and societies. The Social Security Number of a consumer with good credit, for instance, can sell for between $60 and $80 [20]. For example, a web site lists data sale prices for the Brazilian underground: a set of business application account credentials $155–193, a set of credit card credentials: $35–135, a set of online service account credentials: $19; a list of mobile phone numbers: $290–$1,236, and a list of land-line phone numbers: $317–1,931 [17]. Same consumer records may be sold an unlimited number of times by copying. The unlimited sale potential and a large number of consumer records can generate very large rewards. For example, the T-mobile’s breach lost 45,000,000 consumer records. If the consumer records are sold at the lowest cost of just $10 and only about 10% of the records are sought by buyers, it would generate about $45 million. For leaked 864,500,000 records, the hackers could get $172.9 million, assuming that the records are sold at $20 per record one time, and only 0.1% of consumer records are salable.

The sale duration of consumer data depends on use purposes. Some of the data are used for marketing and others are used in committing identity fraud. Most personal data such as phone numbers and addresses do not change at high frequencies, they may be sold five, ten, even twenty times. Other data such as Society Numbers, Birth day, personal working histories, education details, most biometric data cannot be changed at all, they would be sold for the consumer life times or even sold after the consumer’s death. In most cases, consumer data are sold without knowing the consumer’s age. It is very possible that data brokers may validate consumer’s identities by comparing currently acquired data with those that they collected ten or twenty years ago. Consumer medical records may be sold at the higher price of $200 [20]. Medical records of people with chronic diseases are pursued by hospitals, medical practitioners, and drug companies. If consumer data are used in committing fraud and identity theft, their utilities are perpetual. Personal identity data can generate value even after the consumer’s death. After a consumer lives a life, he must leave various vestiges. Using the identities of diseased persons is very common because the deceased persons cannot defend against the fraud. Naturally damages will fall on governments, societies, non-internet companies, and insurance companies. Due to unlimited sales times, hackers or those who trade consumer data would get very large rewards. It is obvious that most personal privacy data do not produce immediate value to hackers, but may become useful when various pieces of personal data are combined to achieve a sufficiently large scope of personal information. Not all consumer records can produce rewards in all times. When some of them generate wealth, reputation or political standing later, they will become subjects of identity theft. Full damages may be realized any time in each victim’s life time. Damages may befall consumers, governments, non-internet companies, and insurance companies.

When the potential reward is so large, a big internet server must invite a large number of potential hackers. For a website like Facebook, I safely assume that thousands of hackers (M) may routinely attack it. I assume that the reported mass data breaches may reflect only a few events that have been detected. There might be long delays between the first action of attacks and eventual reported data breach. It is possible that real problems may be brewed when there is absolutely no sign of any problem. This is very similar to a cancer, which is being made years or decades before a tumor is detectable. I suspect that sound judgment of data security cannot be based on the present performance or lack of sign of data breaches, but must on the totality of systematic designs, use characters, and all parameters concerning the incentives (which are defined by N, T and M). If a system is under attack of M hackers for rewards multiplied by N and T, any claim of data security can be defeated by insider attacks.

In an internet server, a successful hacking of one account is same as successful hacking of all accounts (N). Moreover, each consumer record may be sold many times (T). The enhanced rewards come from those two multipliers N and T. If the vendor has only one account which would be sold at $40, the potential reward is $40*N*T. If the consumer record can be sold for 20 times, the hacker can generate $800. If the system hosts 10 million consumer records, a successful hacking could generate $8 billion rewards for the population's life assuming that all records are good. When hacking can generate so large rewards, they can secure insiders' assistance by offering sufficiently high rewards. It is clear the prevalent data breaches can be stopped only by reducing the number N of consumer accounts and reduce sales number T. Early computer security model was evolved without considering the impacts of the two risk multipliers N and T.

The second incentive is profits from selling consumer data by data brokers. It is anticipated that some hacked consumer records may be laundered as legitimate consumer data in the data broker market. For example, a hacker may hold himself out as a data broker, and sells hacked data to other data brokers. In the chain of data trading transactions, the origins of consumer data are lost. I must suspect that some of the 45 million data records leaked from t-Mobile and 20 million exposed from Apple may become sources of consumer data in the data broker market. They may be sold, traded, or used for marketing perpetually.

Those computations imply that a workable solution to insider-assisted attack are reducing the number of account (N) or using different account access methods, restricting consumer data trade, and imposition of full liability for all damages from data breaches. Those changes will jointly reduce hacker number (M).

C. Reduce Three Parameters N, T and M

Reducing consumer account number N is not an acceptable option for any business. Then, an alternative measure is reducing the amount of information in the consumer data or add an unbreakable huddles to unauthorized access. The e-commerce is based on the population model and the old data security model. By following the population model, all users are treated in the same way as far as security is concerned. By following the data security model and business trust notion, the vendors can be fully trusted and consumers' account data can be stored in plain text so that any employees of the vendor can read, copy and use them. In this model, the vendors not only have the right to charge fees to the consumers for their services, but own consumer data and thus can sell them like their properties. The vendor then treat all consumer data in the same way by applying the population model, resulting in the convention that all consumer data are protected in the same way. Since the server must validate consumer's identity, the password must be stored on the server or must be derivable from other information stored on the server. Application of business trust and the population model thus result in N times of risk of data breaches. If any person knows security measures, password location, and key stretching method, the person can easily crack the security measures. To get a better chance of success, hackers may seek help from insiders by providing sufficient monetary rewards.

To get rid of risk of account multiplicity, an effective security measure is encrypting each consumer's data using an encryption key provided by the consumer and storing the encrypted data on the server, and whenever the vendor needs to use the data, the consumer provides the encryption key that is used to decrypt the encrypted data to produce usable data. The vendor uses the usable data once to complete the business transaction. The encryption key cannot be derived from stored information on the server, and thus insiders could not offer help. Now, the potential reward will decrease by N folds. The measure will also eliminate the T because when consumer data are limited, they lost market value. So, they have lower or no resale values. When reward is so low, fewer of hackers will stay in the hacking business. This will reduce M. The measure can reduce N, T and M to the minimum, hacking will become a trade of the past. In addition, a rule is adopted to use the minimum consumer data for conducting business transactions. This view is also suggested by other scholars [19].

The new data security model is based on cost and reward analysis. I will consider the cost for hacking one single account by brute-forcing encrypted data. The average cost of electricity in the US is $0.12 per kWh. For a single server, it would need about 3741 kWh annually. That would be about $450 per year for one machine. If a machine can do 10^{14} decryption attempts per second, the machine could do about $365*24*3600*10^{14}=3.15\times10^{21}$ decryption attempts per year. The hacker would need to do 2^{255} decryption operations on average to search the half of the key space so that it would need $2^{255}/(3.15\times10^{21})\approx1.84\times10^{55}$ machines. The total cost would be $\$450\times1.84\times10^{55}=\$8\times10^{57}$. Brute-forcing a 256-bit key would cost about 10^{44} times of the Gross World Product (GWP) or $\$63\times10^{12}$ [21]. This cost does not include the costs of depression of 1.84×10^{55} machines in one year and labor costs. Similarly, cracking a symmetric 256-bit key used by AES by brute force would require a search of 2^{256} key space. The effort would require 3.67×10^{55} years [22], but the average success time would be 1.83×10^{55}. In both computations, estimated costs do not include additional expenses that might arise from operational errors, machine failure, comparison time, and costs of human labors. The actual costs and time would be much more. All predictions strongly imply that brute-force attacks for a 256-bit key is unfeasible [22].

Even if encryption is using a 64-bit encryption algorithm, the total time to search all keys by using a standalone server would be $t=2^{64}/(3600*10^{14})\approx51$ hours. Based on personal experience, a one-trip internet interaction would take 1 second, depending on internet speed, server speed, server load or total connections, and the size of rendering page. By brute-forcing over an internet connection, searching a 32-bit key would take a total time: $t=2^{32}/3600\approx1.2$ million hours or 137 years. Those computations imply that brute-forcing over normal internet connections is unfeasible even for a 32-bit encryption algorithm. There are all kinds of other huddles such as connection time limits, attempt limits, AL detection, etc.

Those computations imply that if consumer data are encrypted by a 128-bit or better algorithm at the consumer machine and then uploaded to the vendor server, the data cannot be cracked by any method. The data are immune from insider attacks and legal process.

III. A NEW DATA SECURITY MODEL

A. Use Minimum Data In Conducting Business Transactions

I will explore what kind of consumer data are required for e-commerce. A start point is examining cash transactions. In a cash transaction, only things that must be trusted by two sides are currency authenticity and product quantity and quality. Such transactions do not require anything such as personal address, identity, phone number, email address, and street addresses. There is no need to collect any of those items. This model does not work well in two kinds of situations. In one kind of situation, the products in exchange are very complex and thus cannot be determined by an on-site inspection. Thus, what is required is more time for the buyer. In another kind of transactions, the seller will sell products or services on monthly basis and the seller needs to secure that timely payment will be made in given time intervals, like an utility contract between a consumer and a utility company. In this kind of arrangements, the seller must ensure that buyer will make timely payment each time. To enforce each payment, the seller needs to know the person and thus there is a need to collect information on the personal address, phone number, birthday, and payment mode. Those information is merely incidental to the need to enforce the agreement or secure the last payment in the event that the arrangement ends. Most studies on trust are revolved around vendors [23-25], some measures such as third party assurance [24] and legal remedies [25] would be similarly used as measures against consumers. Indeed, the most usable measures are evaluating consumer's credit records and demanding a reasonable deposit for the consumers with poor credit ratings. Most personal information can be avoided if a proper arrangement can be made.

B. Detailed Technologies Based on the Oldest "Do-Not Tell" Principle

The new technologies are based on the oldest principle that the safest measure is not sharing one's private information with anyone. The technologies to embody this principle can be described as follows:

1. Upload personal data in an encrypted form to a server and store the data in an encrypted form on the server. This is a one-time effort.
2. Send an authorization to conduct the business transaction such as a making a payment. It normally takes a few seconds to complete.
3. Server responds with an acknowledgment.

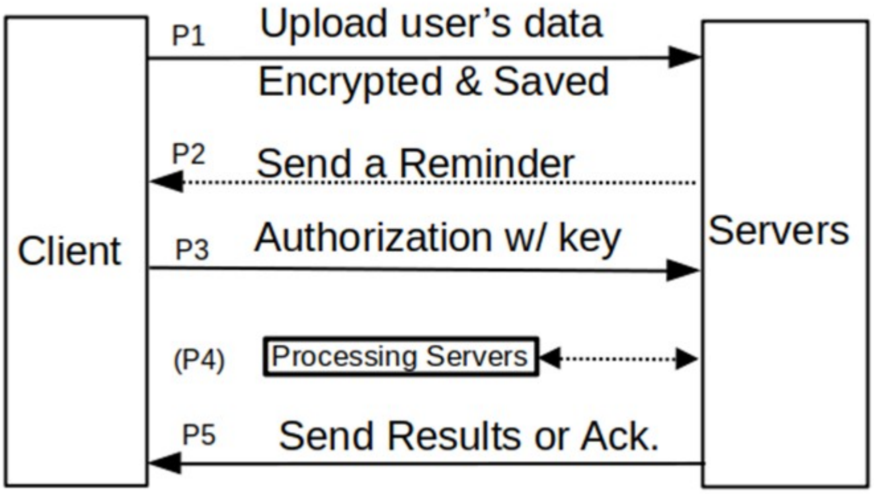
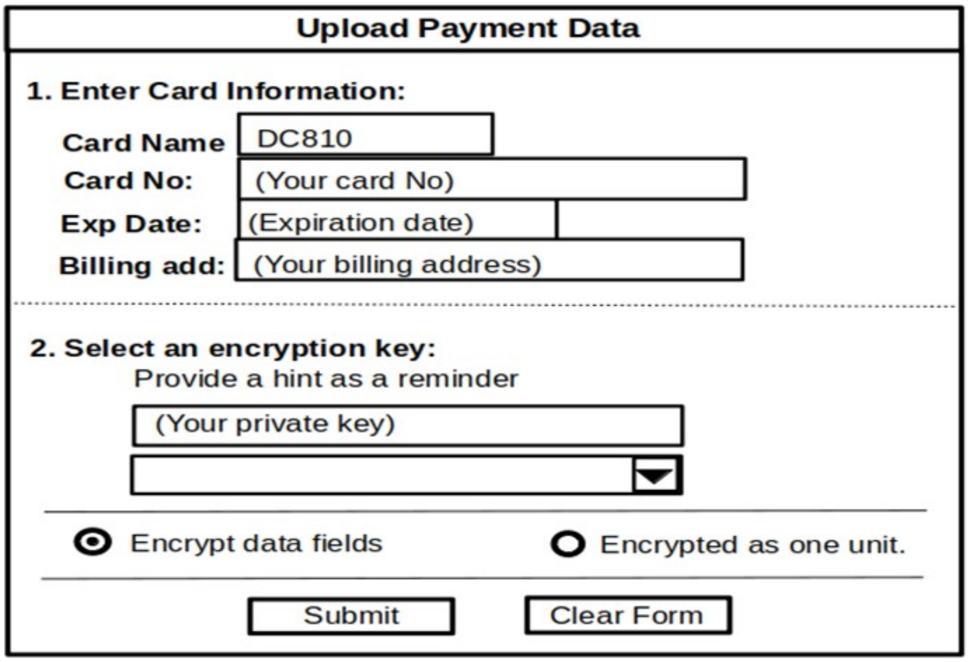


FIG. 1. The Process For Processing A Payment Using Encrypted Consumer Data.

For example, a payment can be completed by using the following exemplar user interfaces (FIGS. 2-4):



Upload Payment Data

1. Enter Card Information:

Card Name DC810
Card No: (Your card No)
Exp Date: (Expiration date)
Billing add: (Your billing address)

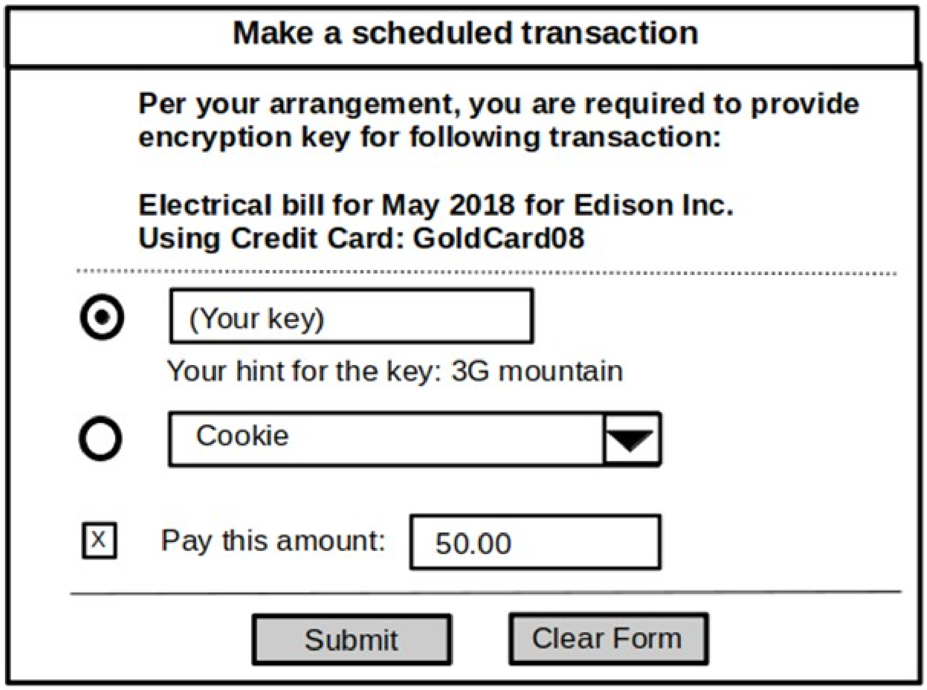
2. Select an encryption key:
Provide a hint as a reminder
(Your private key)

● Encrypt data fields ○ Encrypted as one unit.

FIG. 2. An Exemplar User Interface for Uploading Payment Data.

All sensitive information, as shown in FIG. 2, is encrypted on a client computer or a server by using a user-provided encryption key and stored in encrypted form on the server. The consumer data are not usable by the server without the user-provided key. The consumer can make a payment, which may be initiated by the server.

In making the payment, the consumer or the user is required to provide only an encryption key or the location of the key, and the amount (if the user wants to override the amount in the setup), as shown in FIG. 3. The server gets the encryption key and uses it to decrypt the stored payment data to produce usable data and process the payment on this server or a connected server. The server then provides an acknowledge message as shown in FIG. 4.



Make a scheduled transaction

Per your arrangement, you are required to provide encryption key for following transaction:

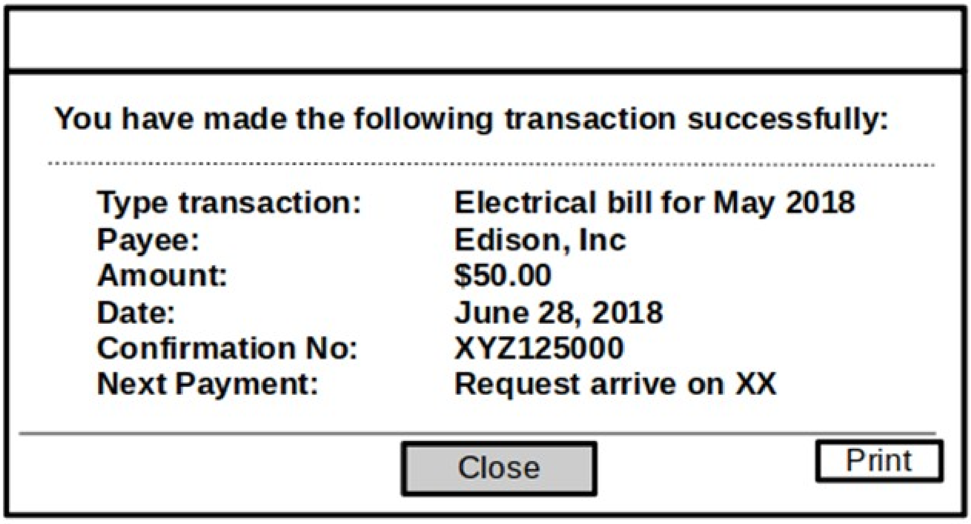
Electrical bill for May 2018 for Edison Inc.
Using Credit Card: GoldCard08

● (Your key)
Your hint for the key: 3G mountain

○ Cookie

☒ Pay this amount: 50.00

FIG. 3. A User Interface for Authorizing a Payment.



You have made the following transaction successfully:

Type transaction:	Electrical bill for May 2018
Payee:	Edison, Inc
Amount:	$50.00
Date:	June 28, 2018
Confirmation No:	XYZ125000
Next Payment:	Request arrive on XX

FIG. 4. An Exemplar User Interface Showing An Acknowledge of Payment.

All sensitive data on the server are encrypted by using different keys and potentially using different encryption algorithms. Hackers will not brute-force consumer data from an account for about $50 to $200 potential return per success or $800 for the person's lifetime. To hackers, all data from the server's database is useless. This concept will lead to a new data security ecosystem: insiders-assisting hacking will be out of business.

C. Feasibility of the Technologies

In conducting a business transaction, personal data, credit card data, and most confidential data are not the type that can lose permanently. If the encryption key is lost, the maximum burden is re-creating the transaction data. The encryption key does not require very high strength. A key like a street name, an object name, animal name, or any combination is enough to defeat the incentive of cracking the encryption key because the hacker is not in a position to know anything about the consumer and other consumers. Such keys can be effective even if they are grossly insufficient for other security purposes. Due to impossibility to brute-force an encryption key in practice, the technologies can achieve the benefits of the "do-not-tell" principle that will never become obsolete. The new technologies can put an end to future consumer data breaches, which were created, promoted and proliferated by the flawed population model. If encryption and decryption are always performed at the client computer, there are very few options for the hackers. The technologies have following advantages:

1. Completely eliminate the incentives for hackers to crack an online server system's security measures to get millions of account data for unlimited sales. On the server, each consumer data record is encrypted by a user-provided key. A hacker cannot successfully crack millions of consumer data records at once.

2. A payment or other transaction can be conducted in seconds, even faster than using two-factor authentication. The time can be even shorter than what is needed to make a regular payment. For other transactions, the time is just as short as typing or pointing a location of the encryption key.

3. Increase consumers' confidence in using debit cards, banking cards, visa cards, master cards, American express cards, etc. in all kinds of commercial environments. When consumers are finally assured that their personal data and financial information cannot be hacked, they will be more willing to arrange such transactions.

4. Reduce the burden on the businesses to secure customer data online; thus, small and middle-sized businesses can accept and use credit cards, banking cards, personal information, etc. without exposing their liability. As long as the consumer data are encrypted by 128-bit or higher, the business owners do not need to worry about data security. The responsibility to control consumer information is given back to consumers. There is no need to get this hot potato.

5. Reduce the burden on consumers in learning new security technologies, changing account passwords, and updating software. In the past, each new technology can improve data security for a short time. When hackers figure out how to crack the new technology, the business owner has to seek further new technologies. Each new technology will require the consumer to learn new things. At times, they will get into trouble due to their own mistakes in using new technologies. Nearly all consumers have been tired of the never-ending security upgrades that can never end the nightmare of mass data breaches.

6. Reduce the chances of access interruptions caused by consumers' mistakes, hardware incompatibility and software incompatibility, etc. which can be caused by or traceable to newly introduced security measures. After consumers have used the same systems for long times, they have learned how to address certain interruptions. When they run into a new problem with an upgraded security measure, they may have no time to deal with the problems and can be negatively affected in special circumstances.

7. Reduce downtime that can be traced to complex and never-working security technologies and help employees to reduce the burden of navigating multiple layers of security measures that never work. In some working environments, employees may be required to pass two to four layers of login, hardware authentication, etc. None of them can keep out real hackers. Most security measures just kill productivity and cause their clients to spend more money.

8. Provide the best protection for "consumer with lost identity." After a huge number of consumer data breaches, many U.S. residents have become public persons, whose social security number, birthday, address, and other personal data have been stolen in one or more data breaches and those identity data are presumably available for sales in the black market. Most of them have not been hit by fraud or criminal activities because they are not in the circumstance of generating values yet. One big problem is that they cannot have their personal identity protected anymore. By using the new technologies, the consumer can create a new personal identify information. Also, when authorization can be made only by using a user-provided key, hackers cannot bypass. For all those who have lost identity data, their future authentication must be based on their secret keys but not their birth days and social security numbers.

9. Prevent more pipe-line disasters from misuse of personal biometric data such as DNA, blood type, fingerprints, face looks, etc. None of those insider-assisted data breaches can be stopped by using personal biometric data. Since biometric data must be stored on a server which can be accessed from the internet, they can be stolen. After biometric data of consumers on a server are stolen, they will be copied, passed from person to person, and are available for sales in the black market. Under current security technologies, victims cannot protect their identities any more because they cannot change their biometric data. Any claim that biometric data can improve data security is a wish. The new technologies will give the lost-identity persons a second chance to protect their identities and financial data in a different way.

The new technologies help protect the environment and promote sustainable economy. They can stop massive wasteful human activity cycles that start with leaking consumer data, fixing all kinds of problems attributable to damaged personal credits, lost money and properties, damaged reputation and damaged personal health, restoring personal credits, recovering financial loss, restoring personal health, catching and prosecuting thieves, and jailing thieves. All massive activities that resolve around leaked consumer data records do not produce any net benefits to victims and societies. All of those activities damage victims, society, governments, and planet.

The new data security model alters the internet convention that the business has a total control over consumer data. To make this model workable, the rights and obligations on managing consumer data must be modified by agreements, common understandings, regulations, trade regulations, statutes, state laws, local laws, etc. The vendor may use consumer data only when there is a need. It does not own consumer data. After use of the user's data, the usable data is deleted or discarded. The data in the usable form may be saved on the server only for a limited time or just sufficiently long for completing an intended business transaction; and the consumer is obligated to provide an encryption key to decrypt stored data for each transaction per their agreements. In a payment system, the vendor may use a gateway, a payment processor, a central computer connected to a bankcard network or automatic clearing house, etc. Each of involved parties follows the same rule.

D. Impacts of Possible Consumer's Errors

A common online fraud is caused by consumer's errors in response to an fake email. For example, a consumer might receive an email from a bank that urges the consumer to approve a pending transaction. If he did not ascertain the authenticity of the email, the consumer might try to log into the account. In this process, the consumer provides his log-in name and password, but ends up with seeing a message to log back next day. If the consumer did not know that he was responding to a fake website, his credentials information will be used to access his banking account shortly. By using the new security model, regular transactions are sending an authorization for conducting a business transaction such as making a payment. The user interface used in this step of transaction contains only an encryption key for data storage security. If the consumer has responded to a fake payment reminder, the consumer will compromise only the encryption key. Even if the key is acquired by a hacker, the consumer just loses the benefits of data encryption in storage. The key itself has no value. It can be changed in a few minutes. Moreover, even if a good number of people have such compromised keys, the net effects are like a few percents of accounts data are stored in plain text. All normal password protection and data transmission protection are still in force. A few incidental errors cannot return the data ecosystem to the old one.

Another issue is what is the impact of forgetting an encryption key on consumer and businesses. First, consumer data are not the kind of data that can be lost forever. The data are not for permanent storage in nearly all transaction environments. If a consumer has lost his key, he could re-submit his personal information for encryption. The scope of impacts is only on the sensitive information that the user wants to protect. Thus, the burden is very limited. Because the encryption key is not a real privacy data, it has no value in the data broker market. The key can be stored on a personal computer, a cellular phone, and/or recorded in personal books. Its main purpose is creating a data ecosystem that turns hacking into a money-losing trade. Since most hackers are remote from the consumers in most situations, it is impossible for the hackers to know about anything about the consumers. Even if a consumer selects an encryption key based on his birth day (even though a bad idea), home street name, school name, car model, etc, it is not what the hacker knows. The hacker faces the reality he must brute-force each consumer record with no useful hint. An expected maximum reward in the amount of $800 is a lost deal.

E. Application Limitations of the Technologies

The technologies cannot be used to protect data originated from the server. They cannot be used to protect bank transaction data that are generated from the banks. For example, if the account of a bank (e.g, a credit issuer) contains both a bank's credit account and a local bank's account which is used to make monthly payment, the new technologies can be used to protect the bank account of the local bank and personal identity information, but not the credit account of the issue bank. Similarly, the technologies cannot be used to protect vendors' sale records such as items, charges, sale dates, etc. Fortunately, those data are not the kind of data that hackers can use to generate money. The utility of the technologies is based on the assumption that the server is controlled by the business vendor so that tempering of server software is not allowed.

The technologies can be used to protect consumer data for utilities companies such as gas companies, electricity companies, water companies, internet companies, retailers, wholesales stories, chain stories, banks and financial institutes, drug stories, credit report agencies, hospitals and health care providers, insurance companies, technical companies, large restaurants, chain restaurants, airlines, bus companies, transits, metro, schools and universities, government agencies, and any other vendors that have needs to keep credit cards, banking cards, and personal data for convenience of the companies. If the technologies are used, banks reduce financial loss from credit card frauds, vendors avoid liabilities for losing personal and financial information, consumers save time on aftermath effort to restore personal privacy and changing passwords, consumers do not need to keep learning new security technologies that do not work for a long time, and employers can see improved work productivity due to reduced interruptions by failed security measures. The technologies may be used to protect staff's personal data and sensitive official data if security responsibilities can be allocated to specific persons.

When the technologies are deployed, consumers should be reminded that the vendor cannot recover lost encryption key. If a consumer wants to keep a document or information as the sole source of storage, the consumer must keep the encryption key permanently. Due to separation between the consumer and the hacker in space, the encryption keys do no need to have high security strength. If consumer data must be protected permanently, the consumer should keep the key in at least two other sources. To avoid that the key is sold by employees of key keepers, the use purpose of the key should not be disclosed to anyone.

F. Existing Technologies Failed to Prevent Mass Data Breaches

All existing technologies are based on both the old data security model and the population model, and thus must have three vulnerabilities: similar or unique exposures of data risk, risks caused by insider attacks and insider collusion, and exposure caused by abusive legal process. The real reward is about N*T*Selling price. The two multipliers (N and T) and current system designs are the main reasons for failure of all existing security measures. Such internet server must be the target of attacks by M hackers. All existing technologies could not protect data storage security. All authentication protocols such as multiple-factor validation [27] and email or phone code and new technologies are ineffective against insider-assisting hacking and abusive subpoenas. If the hacker can buy insider help, they can bypass whatever security measures. One possibility is taking the user password database and even whole the consumer account database from back-end. Another possibility is exploring software design features and possibly planting malware to bypass or disable server security features so that the hacker can get account data or passwords. Afterwords, the hacker can have years to access and acquire account data, most probably by automatic algorithms. For a large system like Facebook, a large number of employees and consultants may know its software structure, system components and calling convention of all programs, and security measures. It is even possible that system developers may leave attack points in software. When rewards are so high, an employee, who may help a vendor to improve security in his official capacity, may develop a scheme to attack the same system for hackers. For the same reason, security tokens [28] cannot protect against insider acts and abusive legal subpoena. When thieves cannot get secured properties, there is a chance that the thieves will stalk and assault the property owner to gain access. If the item is secured with a biometric device [29], the damage to the owner could cost more than the secured property. For example, in 2005, Malaysian car thieves cut off a man's finger when attempting to steal his Mercedes-Benz S-Class [30]. If biometric data are stored on a server, hacking and acquiring the biometric database from the server can enable thieves to create artificial fingers, faces, etc. to gain access in target attacks. As long as the server has copies of consumer data to be used for comparison, a hacker must be able to get the data for later target attacks. Most recent technologies are AI in Cybersecurity [31]. AL may help organizations to detect, predict and respond to cyberthreats in real time using machine and deep learning. While they may mitigate attack risks, its capability would depend on design of AL. Those knowing AL features must be able to find ways to avoid their detection. I predict that they might cause unintended problems and unwanted interruptions. The biggest future risks are acquisition of personal biometric data. Steeling of biometric data will turn their owners into a "public" figures who have lost personal biometric secrets. I predict that no data security measure cannot be cracked if insiders are involved. Each mass data breach can have severe consequences because it negatively affects business operations, stock prices, customer trust, regulatory actions, and liabilities to different persons and parties [32-42].

Encryption of consumer data by using a consumer's key by using a high strength encryption algorithms will have a similar protection of "do-not-tell", absolutely the highest security measure.

IV. CONCLUSION

The frequent mass data breaches can be traced to the influence of the old data security model and population-based approach. In such a system, successful cracking of one account is naturally rewarded by expected sale price multiplied by account number N and selling times T. The potential rewards to hackers for each successful cracking of a business server are so large that no technologies can ever prevent insider-assisting attacks. The extremely large rewards must invite cyber attacks from a large number (M) of hackers. Past data security technologies have not addressed the incentives, perpetual impacts of leaked consumer data on consumers, governments and societies. It is fair to find that damages from most large data breaches are shifted to consumers, governments, insurance companies, and non-internet corporations. Identity theft is becoming a national crisis. A workable solution to this problem must be drastic. The new technologies, which comprise encrypting and storing encrypted data and using consumer data upon user-provided key, can change data security ecosystem. They can eliminate hacking incentives and provide a new protection to the U.S. residents who have lost their identities data.

REFERENCES

[1]. Widipedia. Computer security. Assessed from https://en.wikipedia.org/wiki/Computer_security

[2]. Pennington R, Wilcox D. The Role of System Trust in Business to Consumer Transactions. Journal of MIS. 2004;20(3):197-226.

[3]. Das TK, Tetig B. Between trust and control: Developing confidence in partner cooperation in alliances. Academy of Management Review. 1998;23(3): 491-512.

[4]. Hart PJ, Saunders CS. Emerging electronic partnerships: Antecedents and dimensions of EDI use from the supplier's perspective. Journal of Management Information Systems. Spring 1998;74(4):87-111.

[5]. Hawes JM, Mast KE, Swan JE. Trust earning perceptions of sellers and buyers. Journal of Personal Selling and Sates Management. Spring 1989;9(1):1-8.

[6]. Gefen D. Reflections on the dimensions of trust and trustworthiness among online consumers. DATABASF for Advances in Information Systems. 2002;33(3):38-16. Doi: 10.1145/569905.569910

[7]. Doney PM, Cannon JP. An examination of the nature of trust in buyer-seller relationships. Journal of Marketing. 1997;61(2):35-51.

[8]. Ba S, Pavlou PA. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. MIS Quarterly, 2002;26(3):243-268.

[9]. Chiravuri A, Nazareth D. Consumer trust in electronic commerce: An alternative framework using technology acceptance. In D. Strong and D. Straub (eds.), Proceedings of the Seventh Americas Conference on Information Systems. Atlanta: AIS. 2001, pp. 781-783.

[10]. Hoffman DL, Novak TP, Peralta M. Building consumer trust online. Communications of the ACM. 1999;42(4):80-85.

[11]. Jarvenpaa SL, Tractinsky N, Vitale M. Consumer trust in an Internet store. Information Technology and Management, I. 2000;1(2):45-7.

[12]. Lee MKO, Turban E. A trust model for consumer internet shopping. International Journal of Electronic Commerce. Fall 2001;6(1):75-91.

[13]. Wikipedia. List of data breaches. Accessed from https://en.wikipedia.org/wiki/List_of_data_breaches

[14]. Wikipedia. Data breaches. Accessed from https://en.wikipedia.org/wiki/List_of_data_breaches

[15]. Privacy Rights Clearinghouse. Data Breaches (database). Accessed from https://privacyrights.org/data-breaches

[16]. Song V. (January 17, 2019) Mother of All Breaches Exposes 773 Million Emails, 21 Million Passwords. Gizmodo. Retrieved from https://gizmodo.com/mother-of-all-breaches-exposes-773-million-emails-21-m-1831833456

[17]. Trendmicro. The Global Black Market Prices: These are some of your personal data and their corresponding prices: https://www.trendmicro.com/vinfo/us/security/special-report/cybercriminal-underground-economy-series/global-black-market-for-stolen-data/#section-2

[18]. Tucker C, Neumann N. (May 1, 2020) Buying Consumer Data? Tread Carefully. Analytics And Data Science Harvard Business Review. Accessed from https://hbr.org/2020/05/buying-consumer-data-tread-carefully

[19]. Martin B. The Unregulated Underground Market for Your Data: Providing Adequate Protections for Consumer Privacy in the Modern Era. 105 Iowa L. Rev. 865 (2020).

[20]. Max E. How Companies Turn Your Data into Money, PCMAG (Oct. 10, 2018, 8:00 AM), https://www.pcmag.com/article/364152/how-companies-turn-your-data-into-money [https://perma.cc/Y8SN-TS53].

[21]. Stockexchange.com How much would it cost in U.S. dollars to brute-force a 256-bit key in a year? https://crypto.stackexchange.com/questions/1145/how-much-would-it-cost-in-u-s-dollars-to-brute-force-a-256-bit-key-in-a-year

[22]. Wikipedia. Brute-force attack. https://en.wikipedia.org/wiki/Brute-force_attack#Theoretical_limits

[23]. Hart PJ, Saunders CS. Power and trust: Critical factors in the adoption and use of electronic data interchange. Organization Science, 1997:8(1):23-42.

[24]. Kimery KM, McCord J. Third-party assurances: mapping the road to trust in e-retailing. Journal of Information Technology Theory and Application, 4. 2002;4(2):63-82.

[25]. Sitkin SB, Roth NL. Explaining the limited effectiveness of legalistic "remedies" for trust/distrust. Organization Science, 1993;4(3):367-392. (legalistic remedies often have limited effectiveness in addressing trust problems in organizations)

[26]. Schurr PH, Ozanne JL. Influences on exchange processes: Buyers' preconceptions of a seller's trustworthiness and bargaining toughness. Journal of Consumer Research. 1985;11(4):939-953. Doi: 10.1086/209028 (a seller's expected trustworthiness-plus-toughness in bargaining led to high levels of buyer–seller cooperation and agreement and a higher level of buyer concessions.)

[27]. Wikipedia. Multi-factor authentication. Accessed from https://en.wikipedia.org/wiki/Multi-factor_authentication

[28]. Wikipedia. Security token. Accessed from https://en.wikipedia.org/wiki/Security_token

[29]. Wikipedia. Biometrics assessed from https://en.wikipedia.org/wiki Biometrics#Danger_to_owners_of_secured_items

[30]. Kent J. (31 March 2005). Malaysia car thieves steal finger. BBC Online. Kuala Lumpur. Archived from the original on 20 November 2010. Retrieved 11 December 2010.

[31]. Capgemini Research Institute. (2019) Reinventing Cybersecurity with Artificial Intelligence: The new frontier in digital security. Assessed from https://www.capgemini.com/wp-content/uploads/2019/07/AI-in-Cybersecurity_Report_20190711_V06.pdf

[32]. IBM Security. Cost of A Data Breach 2020. Assessed from https://www.ibm.com/security/digital-assets/cost-data-breach-report/1Cost%20of%20a%20Data%20Breach%20Report%202020.pdf

[33]. NMER (June 2018). Economic and Financial Consequences of Corporate Cyberattacks. The Digest No. USA. Assessed from https://www.nber.org/digest/jun18/economic-and-financial-consequences-corporate-cyberattacks

[34]. Lease ML, Burke TW. Identity Theft: A Fast-Growing Crime. Journal FBI Law Enforcement Bulletin. August 2000;69(8): 8-13.

[35]. ITRC. 2021 Consumer Aftermath Report: How Identity Crimes Impact Victims, their Families, Friends, and Workplaces. Assessed from https://www.idtheftcenter.org/publication/identity-theft-the-aftermath-study/

[36]. Lawson L. (January 21, 2022) 2021 Identity theft statistics. Trends and statistics about identity theft. Assessed from https://www.consumeraffairs.com/finance/identity-theft-statistics.html

[37]. Newman GR, McNally MM. (July 2005) Identity Theft Literature Review. Unpublished report of U.S. Department of Justice. Award Number: 2005-TO-008. Document No.: 210459.

[38]. Benner J. Mierzwinski E, Givens B. (May 2000). Nowhere to turn: Victims speak out on identity theft. California Public Interest Research Group and the Privacy Rights Clearinghouse. Assessed from http://www.calpirg.org/consumer/privacy/idtheft2000/idtheft2000.pdf

[39]. Gordon GR, Curtis GE. (2000). The growing global threat of economic cyber crime. National Fraud Center, Inc. Assessed from http://www.lexisnexis.com/rissolutions/conference/docs/cyber.pdf

[40]. Insurance Information Institute. Facts + Statistics: Identity theft and cybercrime: The scope of identity theft. Assessed from https://www.iii.org/fact-statistic/facts-statistics-identity-theft-and-cybercrime

[41]. Giact. U.S. Identity Theft: The Stark Reality. Assessed from https://giact.com/identity/us-identity-theft-the-stark-reality-report/

[42]. Weisbaum H. (July 30, 2018, Updated July 30, 2018). The total cost of a data breach — including lost business — keeps growing. Assessed from https://www.nbcnews.com/business/consumer/total-cost-data-breach-including-lost-business-keeps-growing-n895826.