The Impetus to Creativity in Technology

Alan G. Konheim
Professor Emeritus Department of Computer
Science University of California
Santa Barbara, California 93106
konheim@cs.ucsb.edu ✉agkonheim@cox.net

Abstract: We describe the ensuing developments from two now well-known publications in the 20th century. They contained significant and seminal technological results, a paper by Claude Shannon in 1948 and a patent by Horst Feistel in 1971.

Shannon's paper announces the challenge at the start, writing ``the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected *sent+ at another point.` Shannon's Coding Theorem established the relationship between the probability of receiving the message in error and transmission rate measuring the process efficiency. Shannon proved the existence of codes achieving *optimal* performance, but it required forty-five years to exhibit an actual code achieving it. The optimal-efficient Shannon codes are responsible for a wide range of communication technology we enjoy today, from GPS, to the NASA rovers *Spirit a*nd *Opportunity* on Mars, and also to worldwide communication over the Internet.

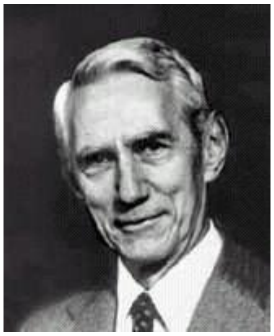
US Patent #3798539A filed by the IBM Corporation in1971 described Horst Feistel's *Block Cipher Cryptographic System, a* new paradigm for encryption systems. It was a departure from the current cryptographic technology based on shift-register stream encryption for voice and the variety of the electro-mechanical cipher machines introduced nearly fifty years before. Horst's vision was its application to secure the privacy of computer files. It was invented at a propitious moment in time and implemented by IBM in *automated teller machines* for the Lloyds *Bank Cashpoint System*. The emergence of E-Commerce in the next decades would far overshadow the value of encryption in banking business. Finally, it has drastically altered the SIGINT mission of the National Security Agency. While Horst Feistel did not directly participate to any of these commercial applications of cryptography, the traditional interpretation of the *law of cause and effect* certainly describes the effect of Feistel's patent.

Key Words: A.A.Albert, AFCRC, ATM, Banking, Block Ciphers, Claude Shannon, Communication, Cryptanalysis, Cryptography, E-Commerce, Horst Feistel, IBM, NSA

1.0 INTRODUCTION

Creativity is a process resulting in something new and potentially valuable. Creativity is hard to quantify and more importantly, difficult to assess its effect on technology. Additionally, realization of the social and economic effects created may not occur until well after their first appearance. In this paper, I discuss two examples of the developments that followed generated by creativity. The first published in 1948 by Claude Shannon, the second by Horst Feistel in 1971. Shannon's paper dealt with the reliability-efficiency constraints in communication, while Feistel proposed a method to protect the secrecy of data communications. Their contributions triggered the *start* of research leading to the current understanding of the relations between the fidelity, efficiency, secrecy and the source-authenticity of data transmission and remarkably involved mathematical techniques understood by both authors.

2.0 SHANNON'S CODING THEOREMS[1]



Claude Elwood Shannon (1916 – 2001) was an American mathematician and electronics engineer. While working for Bell Telephone Laboratories, he published ``A Mathematical Theory of Communication` *55]. The paper determined the maximum rate at communication of messages over a *noisy* transmission medium with bounded error probability could be achieved In a basic framework, a message conveys *n* bits of data $\underline{d} = (d_0, d_1 \ldots d_{n-1})$ over the *memoryless binary symmetric channel* (BSC).Memory-less means the errors for different bits are statistically independent events. When binary data is transmitted over this BSC, each bit is decoded at the receiving end correctly with probability $q = 1-p$ and incorrectly, with probability *p.* Shannon defined the *channel capacity C* as the maximum rate at which reliable data transmission may occur; the formula $C = 1 + p \log 2\ p + q \log 2$ q was derived. The value of *p* depends on the transmission medium (bandwidth), the signal power and the interfering noise according to the *Shannon-Hartley Law*. The assumption $0 \leq p < ½$ may be made without loss of generality. To achieve reliability, Shannon coded the data by supplementing it with redundancy, including *k* check bits $\underline{c} = (c_0, c_1 \ldots c_{k-1})$ to the *n* bits of data $\underline{d}$ to form the *transmitted* message $\underline{x} = (\underline{d}, \underline{c})$. Transmission might corrupt $\underline{x}$ resulting

[1] Photograph courtesy of the IEEE History Center.

in the *received* message $\underline{y}$= $\underline{x}$+ *error*. When $\underline{y}$ is decoded, data $\underline{d}^*$is recovered. The redundant channel coding is viewed as successful when the original data is recovered, meaning $\underline{d}^* = \underline{d}$. Adding redundancy might achieve this at a cost; adding k check bits lowers the *rate R* at which data bits are effectively transmitted to $R = n/ (n+k)$. Shannon proved two propositions; *i*) reliable coding schemes exist to achieve a maximum rate R bounded by C and, *ii*) a converse, if a coding scheme for data of length transmits at a rate above C, then the error probabilityε must be bounded away from 0 as $n \to \propto$.

Before Shannon's paper, it was known that transmitting many *copies* of each data bit, say $2j+1$, the (error) probability that $\underline{d}^* \neq \underline{d}$ could be made smaller than ε for a fixed number of data bits n, provided j was large enough. The maximum likelihood (ML) decoding rule decides for each of the $2j+1$ duplicated transmitted information bits that value 0 or 1, which occurs in the majority. However, this repetition coding reduces the rate from 1 to $R = 1/ (2j+1)$ at which the n bits of data are being delivered. In order to achieve a small decoding error probability ε, the rate R decreases to 0 as $n\uparrow$and $\varepsilon\downarrow$. Shannon proved that communication with arbitrarily small decoding error ε did <u>not</u> necessitate a compensating slow rate. He proved that coding schemes exist which both transmit at a rate close to the channel capacity C and achieve error probability smaller than ε and this was the best possible result.

Shannon's BSC Coding Theorem: For everyε and $\delta > 0$ and data $\underline{d}$ of bit length n which satisfies $n>n^* = n^*(\varepsilon,\delta, p)$ there exists

- a code $\underline{d} \to \underline{x}$ with code words of length m and
- a decoding algorithm achieving an error probability smaller than ε, for which
- the coding rate $R = n/m$ satisfies $C-\delta \leq R < C$

Strong Converse to Shannon 's BSC Coding Theorem: [61] For codes of rate greater than the channel capacity C, the error probability converges to 1 as the data lengths $n \to \propto$. In other words, the maximum achievable reliable transmission rate is C**.**

Shannon proved the existence of coding schema, but did not explicitly show the required redundancy. An implementation would appear 45 years later

Shannon's paper referred to the soon to be published paper [27] authored by his colleague R. W. Hamming. It described a class of *parity check codes* for the BSC. The subject of (error-correcting) *codes* was to occupy the communication community for more than two decades. It began with the study of *block algebraic codes*.

2.1 BLOCK ALGEBRAIC CODES

Hamming investigated codes which were subsets of the additive group $Z_{m,2}$ of 2^m (m-vectors) sequences of 0's and 1's of length $m = n + k$ with $m = 2^s - 1$, $n = 2^s – s - 1$ and $k = s$. The *weight* of an m-vector $\underline{x}$ is the number of 1's; the *Hamming distance* between two m-vectors $\underline{x}_1$ and $\underline{x}_2$ is the weight of their sum $\underline{x}_1 + \underline{x}_2$. The sphere of radius r about an m-vector $\underline{x}$ consists of all m-vectors $\underline{y}$ at distance at most r. A Hamming code word $\underline{x} = (\underline{d}, \underline{c})$ in $Z_{m,2}$ with $m = n+k$ results by transmitting along with the n data bits $\underline{d} = (d_0, d_1 \ldots d_{n-1})$, k parity check bits $\underline{c} = (c_0, c_1 \ldots c_{k-1})$.

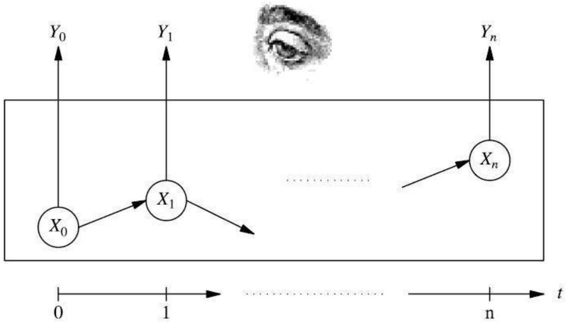
An *r-error-correcting code* is a subset of **C** of $Z_{m,2}$of $2^n m$-vectors whose distinct elements (*code words)* have mutual distances at least as large as $2r+1$. The spheres $\{ S\mathrm{r}[\underline{x}] \}$ of radius r about each distinct code word $\underline{x}$ are disjoint. **C** constitutes an *r-error-correcting* code since maximum likelihood decoding (ML) finds the nearest code word to the received m-vector $\underline{y}$. The number of code words in $S\mathrm{r}[\underline{x}]$ satisfies the *Hamming bound* $|| S\mathrm{r}[\underline{x}] || \leq$ 2k. The Hamming codes with $m = 2^s -1$ are *perfect,* in the sense that these spheres fill $Z_{m,2}$. Numerous technical papers [10, 27, 49, 50] constructing error correcting codes and books [7, 34, 36, 46, 47] appeared in the next decade. A related and more complicated problem is the *decoding* of algebraic codes; *given* the received m-vector $\underline{y}$, *find* the closest code word $\underline{x}$.

Obtaining a code attaining the Shannon rate remained elusive, even with the discovery of many new codes generalizing Hamming. At a 1971 IEEE Communications Society Meeting in St. Petersburg (Florida), a leading algebraic coding authority proclaimed [35] ``coding is dead!" While the accuracy of the obituary might be disputed, it was evident a new approach to coding might be needed.

2.2 PROBABILISTIC DECODING

The new avenue of attack was a modification of Shannon's idea to transmit independent replicated versions of the message so that the decoder would be able to reconstruct the original. It involved several different aspects.

1. <u>*Coding System:* </u>Instead of adjoining parity check bits to data, *convolutional codes* generate check digits via the application of a Boolean polynomial function to a sequence of data bits.
2. <u>*Trellis Representation and Decoding:* </u>A *trellis diagram* represents the state transition diagram of a convolutional code; decoding is a search for the most likely path. Viterbi's paper *59] described the search algorithm using *maximum likelihood* (ML) estimation. It determines the closest code word using the Hamming distance as a metric. It differs from the code word which minimizes the *a posteriori probability* (APP) of the i_{th} data bit d_i,
3. <u>*Soft-decision versus hard-decision decoding:*</u>
    - *hard decision decoding*; if y_j value is in a specified set, then decode it as 1; otherwise, as 0;
    - *Soft decision decoding;* decode the y_j value into say, a real-valued triple with *confidence values.*
4. Hagenauer and Hoehner [26] extended in 1989 the *soft output Viterbi Algorithm* (SOVA) for paths in the trellis using the APP metric.
5. In 1974, Bahl *et al* [5] suggested an entirely different approach to decoding convolutional code output based on the *hidden Markov model* (HMM).



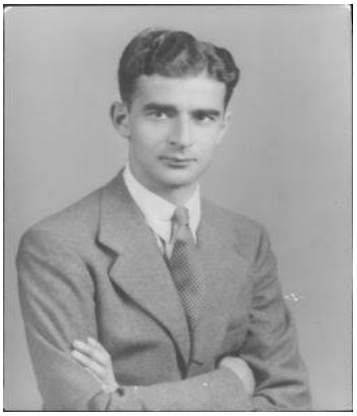
Hidden Markov Model

A HMM supposes the values of the Markov state X_j are not directly observed, but only indirectly through the observable state Y_j. HMMs originated in cryptanalysis [6] in which the plaintext X_j is hidden and the ciphertext Y_j is observable. A paper applying the HMM model in the analysis of problems in a variety other signal detection environments appeared [48].

6. The penultimate step in the process started in 1948 by Shannon occurred in 1993. Berrou *et al* [8] announced a new class of turbo codes which achieves the Shannon channel capacity asymptotically with increasing data bit size n.

Conclusion: Shannon's genius started modern communications theory, but a practical implementation achieving both reliable and efficient communication required significant developments over the next 45 years to reach today's state.

3.0 HORST FEISTEL[2]



Hindenburg Ernst Richard Horst Feistel was the son Richard and Helene Freudenreich Feistel of Frankfurt an der Oder (Germany). He was born in (East) Berlin on January 30, 1915. Some information on Horst is contained in Steven Levy's amusing book Crypto [33]. Feistel joined the Computer Science Department at the IBM Research Center in 1968. I became his manager when he transferred to the Mathematical Sciences Department in 1971; our offices on Aisle 7 were adjacent. Horst Feistel was 19 years older than I, when he joined my group. While we were both cordial and even quite friendly, I did not have the same close personal relationship I still enjoy today with my IBM colleague Roy Adler.[3]

In 1933, Adolph Hitler declared his intention to rearm Germany in clear violation of the Treaty of Versailles; Hitler additionally instituted universal military service two

[2] Photograph circa 1948, Courtesy Mrs. Peggy Feistel Chester.

[3] My favorite Feistel anecdote concerns Horst's IBM work schedule; Horst was an early bird claiming to arrive at 7:00 AM and usually leaving at around 11:00 AM. This schedule was acceptable at IBM Research since he was active and *produced quality research.* My close friend and colleague Roy Adler, to whom the same productivity principle applied, arrived at 11:00 AM. Their paths often crossed in the parking lot at 11:00 AM and the following exchange is alleged to have taken place.

Horst to Roy: Ich wünsche dir einen schönen Tag! (Have a nice day!)
Roy to Horst: Gehabt einen schönen Tag? (Had a nice day?)

This work schedule was fine for IBM Yorktown, but caused problems for Horst at the end of his career. He requested and received a one-year sabbatical at the IBM Scientific Center in Cambridge (MA). The IBM management there did not have the same work schedule tolerance as IBM Research. It was further complicated when Horst announced that he planned *not* to return to Research after his sabbatical ended, but rather to stay at Cambridge until retirement. I was called by his Cambridge manager; as a result of my attendance at three IBM Manager (Finishing) Schools, I followed the appropriate IBM procedures; namely, I took `two breaths' and alerted my manager who took ….

years later. Horst Feistel's maternal Aunt Gertrude lived in Zurich, having married Franz Meyer who was a Swiss citizen and Jewish[4]. Horst's uncle might have learned about Hitler's intention, perhaps through the caustic commentary on Hitler's military intentions appearing in the Swiss press; for example, in the *Neue Zürcher Zeitung*, whose availability in Germany had been forbidden indefinitely. Concerned about his nephew's future, he urgently advised Horst leave Germany for the safety of United States. Horst left from Bremen (Germany) on March 23, 1934 on the S.S. Bremen arriving in the U.S. six days later. He returned to Zurich in August 1934 and enrolled in the Eidgenössische Technische Hochschule (ETH) in Zürich, returning to Massachusetts as a passenger on the SS Champlain leaving on July 22, 1936 from Le Havre. Horst Feistel continued to be in contact with his aunt and uncle who then lived in New York City while at IBM in the late 1970s.

Some details of his life can be gleaned from various genealogical websites. The extensive Feistel family, located throughout the world, maintains their own website Feistel.org. Entries at the website Familysearch.org, sponsored by Church of Jesus Christ of the Latter-Day Saints, show that Horst Feistel

1. Entered the United States in New York arriving on the SS Bremen on March 29, 1934[5].
2. Resided in Ward 6, Cambridge, Cambridge City, and Middlesex, Massachusetts in the 1940 Census.
3. Married Leona M. Feistel (nee Gage) in 1945; Leona was born on January 4, 1924 and died on August 5, 1990.
4. Horst and Leona were the parents of a daughter Peggy, born April 13, 1946.
5. Lived in Mount Kisco (New York) while working at IBM Research Center.
6. Lived at 920 Main Street in Osterville (MA)[6] during 1985.
7. Horst Feistel passed away on November 14, 1990 in Massachusetts.

[4] In January 2015 I finally located Peggy Feistel Chester, the daughter of Horst and Leona Feistel. While aware of her father's work in cryptography, her knowledge of his contribution was limited and she was enthusiastic about learning of his work. She has supplied me with considerable memorabilia , including photographs, documents, letters , which will appear in a subsequent paper on his personal life during the pre-IBM period 1915- 1968.

[5] *Quaere verum* (seek the truth). Additional entry dates into the U.S. are cited. They correspond to trips to the United States from Europe; for example, one trip made on the completion of his studies at the ETH is cited. They will be explained in a subsequent paper.

[6] I believe that Horst and Leona Feistel owned a home in Dennis (MA) on Cape Cod.

3.1 HORST FEISTEL'S WORK BEFORE COMING TO IBM

Some of Horst Feistel's activities after arriving in the United States may be found at from two sources; an author summary in an IEEE paper[7] and a Security Investigation Report applying for a Clearance for his last job before coming to IBM. The last address for Horst Feistel given in the online MIT Alumni Directory is Dennisport, Massachusetts.

As a German citizen, Horst's freedom to travel might have been limited; Steven Levy writes [33, p.39] that Horst was under *house arrest.* Horst enrolled at MIT as an undergraduate; his studies were successful and he was awarded the degree B.S. Physics from MIT in 1937[8]. His studies in Physics continued and he received the degree M.A. Physics from Harvard University in 1942.

A plan for a National Defense Research Council to oversee scientific research directed towards the impending war effort reached President Roosevelt's desk in June 1940. It originated with Vannevar Bush, the Scientific Advisor to the President along with Karl Compton, President of MIT, and James Conant, President of Harvard. Roosevelt quickly approved. Compton headed up the section of the Council overseeing technologies for detection of aircraft and ships, capabilities that were then absent. Compton asked to host the new laboratory at MIT, was initially reluctant because of the perceived campus reluctance. The Committee eventually persuaded Compton and the MIT Radiation Laboratory (*Rad Lab*) was born in the fall of 1940 at MIT. The name chosen to be intentionally deceptive, creating the perception that the laboratory was working on nuclear physics.[9] The Radiation Laboratory contributed to the development of microwave radar technology in support of the war effort during 1940-45. Rad Lab inventions included interrogate-friend-or-foe beacon systems, more often referred to as *identification-friend-or-foe* (IFF).

Hanscom Air Force Base began its existence while the United States was considering its entry into the Second World War. In mid-1942, the Commonwealth of

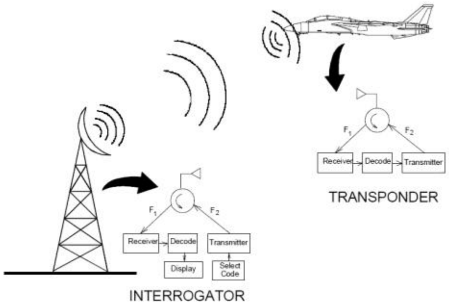
[7] From the author summary in a paper [24] co-authored by Horst Feistel and future IBM colleagues William A.Notz and J. Lynn Smith.

[8] The award of a B.S. (Physics) from MIT in three years while possible seems at least like an unusual event. For example, the four year Bachelor of Science degree in Computer Science at UCSB is rapidly ceasing to be typical. It suggests that Horst was *both* very smart and received credit at MIT for his studies at the ETH in Zurich.

[9] The Manhattan Project to build the bomb would start shortly afterwards.

Massachusetts leased the Bedford airport to the War Department for use by the Army Air Forces. The airfield also served as a test site for research on radar conducted by MIT's Radiation Laboratory[10] and Harvard's Radio Research Laboratory. Hanscom served as a site for testing new radar sets developed by MIT's Radiation Laboratory. World War II established the key military importance of radar. While, the MIT and Harvard wartime laboratories ended in 1945, the Army Air Forces continued some of their programs in radar, radio and electronic research. It recruited scientists and engineers from the MIT's laboratories site at Hanscom Field establishing the Air Force Cambridge Research Laboratories (AFCRL).

While Horst Feistel might have been under *house arrest* in 1936 when we arrived, his fortunes changed in 1944 when he received citizenship, a security clearance and a job. He became a staff member in 1944 at the M.I.T. Radiation Laboratory working on their IFF[11] project. The Mode 4 IFF *challenge X* transmitted to an unidentified aircraft requires a cryptographically determined *valid* response, the encipherment of *X* combined in some way with the aircraft's altitude and identifier.



IFF Mode 4 Challenge/Response Transponder[12]

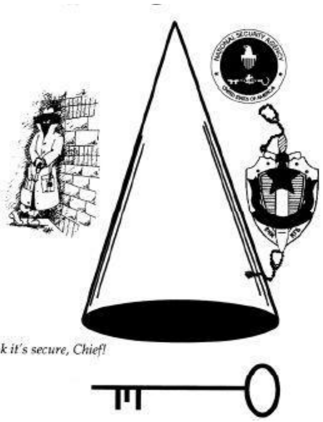
[10] The MIT Radiation Laboratory became one of the largest wartime projects ever, employing nearly 4000 people at its peak. Engineers and scientists from around the country were recruited, and during the war more than 100 different radar systems were developed there.

[11] Identification, friend or foe (IFF) is an identification system designed for command and control. It enables military and national (civilian air traffic control) to identify aircraft, vehicles or forces as friendly and to determine their bearing and range from the interrogator.

[12] This is Figure 1 [page 5.6-2] in Unclassified Report #NAWCWD TP 834, ``Electronic Warfare and Radar Systems Engineering Handbook`, Naval Air Warfare Center Weapons Division Point Mugu, California, April 1, 1997.

Horst Feistel left AFCRC[13]becoming a Member of the MIT Lincoln Laboratory in 1958. Horst Feistel is the author of a 1958 Lincoln Laboratory survey report [21][14] on problems in authenticated communication and control. The report cites the problems of data spoofing and the disruption of communications well as the use of encryption and authentication dependent redundancy, but only in the context of military communications. The acknowledgement in [21] indicates that Feistel concentrated on the security aspect of the data-link project. Before finally joining *Big Blue* in 1968, Horst Feistel worked for the *MITRE Corporation* in 1961.

3.2 I WAS 10 IN 1944 AND UNAVAILABLE TO TEACH CMPSC 178



Reader Cover 👂 UCSB CMPSC 178 (Introduction to Cryptography)[15]

Levy [33, p. 40] writes that ``codes had fascinated him [Feistel] since his early boyhood'; he was now involved in their study and application at AFCRC. While employed there, Horst Feistel may have given access to the classified techniques of cryptographic design, perhaps by the very distinguished American mathematician, A. Adrian Albert[16].

[13] For the first two years of my graduate studies in mathematics at Cornell University (9/1957 -1958), I had a GE Cooperative Fellowship and worked part-time at the *General Electric Advanced Electronics Center,* located at the Ithaca airport. GE had a contract with AFCRC on error-correcting codes at that time. I visited Cambridge meeting Eugene Prange. Horst was gone and we did not meet until 1968 at IBM.

[14] Provided courtesy of Dr. Whitfield Diffie.

[15] The cover of my lecture used when I taught CMPSC 178 at UCSB notes included the *Cone of Silence*, derived from Don Adams 1965 television series Get *Smart,* a spoof of the CIA.I taught CMPSC 178 twenty-one times at UCSB and abbreviated versions in Australia, Hawaii and Israel.

The 2005 book [2] by A$^{3\prime}$s daughter Nancy E. Albert contains her reminiscences of A. A. Albert's work in the cryptographic arena during and after World War II.

What aspects of cryptography might Horst Feistel have learned before coming to IBM? A review of Ms. Albert's book [11] begins with the following statement attributed to David Kahn [29, p. 410]

> Adrian Albert was perhaps the first to observe that, as he put it, `′all of these methods [of cryptography] are very special cases of the so-called algebraic cipher systems.′′[17]

Because of Albert's World War II work in cryptography, it is quite likely that Horst Feistel's exposure to cryptology was significantly meatier than A$^{3\prime}$s November 1941 cryptography lecture. The published version consists of only very basic cryptographic concepts and elementary examples of cipher systems.

Albert was quite influential in shaping Horst's career; Nancy Albert writes that her father was

> ′′… well acquainted with Horst Feistel ′′ and ′′having consulted in private industry and served with some of the leaders of industry on various boards, Adrian knew where the jobs were in mathematics. He advised the young Horst Feistel, who felt his interest in exploring cryptography was being stifled at the National Security Agency, to seek employment at IBM. Feistel later flourished in IBM's *intellectual playground* developing ways to protect privacy in cyberspace.′′

The group theorist Daniel Gorenstein[18], describes the interaction between Albert and Horst Feistel in a paper about the classification of the finite simple groups writing

[16] Abraham Adrian Albert (1905-1972) was a distinguished American mathematician. He received the American Mathematical Society's Cole Prize in Algebra for his work on Riemann matrices in 1939. He was president of the American Mathematical Society during 1965-6. As an applied mathematician, he also did work for the military during and after World War II. One of his most notable achievements was his groundbreaking work on cryptography. The manuscript entitled "Some Mathematical Aspects of Cryptography," was the subject of his invited address at a meeting of the American Mathematical Society in November 22, 1941. It was published and available at the University of Chicago Library [Z103.A33].

[17] Groups, semi-groups, ring, fields and ideals are examples of algebraic systems.

[18] Gorenstein went on to become one of the leading stars in Group Theory. Professor Gorenstein writes in his paper [25]

> ``In the summer of 1957 Horst Feistel's group of cryptanalysts at AFCRC (Air Force Cambridge Research Center) sponsored a research project on classified and unclassified cryptanalytic problems at Bowdoin College. Albert, who had a longstanding interest in cryptanalysis and was a consultant to Feistel's group, invited a distinguished group of university algebraists to participate, including I. N. `Yitz' Herstein[19], Irving Kaplansky, Irwin Kleinfield, Richard Schafer, and George Seligman. In preparation for the project, Feistel's group prepared a long list of classified and unclassified problems. Although many of these were of a field-theoretic or combinatorial character, under Sy Hayden's influence they included a number of purely group-theoretic questions related to a type of cryptographic system then under investigation.`

Some scholars might reason that Horst's cryptographic work was most influenced by the distinguished visitors to AFCRC, including Professor A.M. Gleason. I do not think this is the case as these consultants were *largely* famous for their work in algebra, the focus of the Bowdoin College project. Meeting Horst for the first time, the IBM DES project manager Walt Tuchman asked him the origin of the `LUCIFER` idea. Horst Feistel is quoted by Levy [33, p. 45] as responding, ``The Shannon [*secrecy* [56]] paper reveals all." It is not inconceivable that this publication alone provided the real inspiration for a person of Horst's creative genius. Whatever was the source of the stimulus, Horst Feistel clearly set off in his own direction and offered cryptography a fresh point of view. He provided a clear example of, what today is fashionably championed as, *thinking out of the* [group theoretic and algebraic (cipher) systems] *box*. Horst Feistel certainly derived something from his interaction with each of the consultants; in this sense, they all functioned as his cryptographic mentors.

``The first time I ever heard the idea of classifying finite simple groups was in the mid-1950s in the parking lot of the Air Force Cambridge Research Center at Hanscom Field in Bedford, Massachusetts. After work one day, Seymour Hayden, a member of a group of mathematicians with whom I was a consultant on the design and analysis of cryptographic systems, was describing the thesis problem that Richard Brauer had given him... group-theoretic questions related to a type of cryptographic system then under investigation."

[19] I might have missed joining the early crypto-boat; during 1957-60, while I was a graduate student at Cornell University, Professors Gorenstein, Walter Feit and Itz Hertstein in the Mathematics Department. Professor Hertstein was a member of my examining committee. I chose *without regret* to study under the supervision of Professor Mark Kac in probability theory. I could not have been entirely deficient in algebra, because I passed the Oral Qualifying Examination,

3.3 SHANNON'S SECRECY CONCEPT AND ITS BUILDING BLOCKS

Any possible ideas Albert conveyed to Feistel in1 1944, took place in a classified environment at the AFCRC. It might have included (classified) principles of the design of cryptographic algorithms, extending those appearing in Shannon's 1949`secrecy' paper [56][20]. Shannon's begins by remarking about the similarity between the *reliable* and *secret* aspects of communication.

> ``The problems of cryptography and secrecy systems furnish an interesting application of communication theory. In this paper a theory of secrecy systems is developed. The approach is on a theoretical level and is intended to complement the treatment found in standard works on cryptography [Note: A reference of Helen Fouché Gaines' book ``Elementary Cryptanalysis` is given]. There, a detailed study is made of the many standard types of codes and ciphers, and of the ways of breaking them. We will be more concerned with the general mathematical structure and properties of secrecy systems.`

Shannon defined *perfect secrecy* of a cryptographic system by requiring the *a priori* probability $P_{a\ priori}(M)$ of the a message M be the same as the *a posteriori* probability $P_{a\ posteriori}(M/C)$ after intercept of the ciphertext C.[21] In other words, interception of C does not improve knowledge of the original plaintext message.

An *ideal system* is one in which $H_C(K)$ and $H_C(M)$ do not approach zero as the message length increases $n \to \propto$, where H denotes *entropy;* M is message, K is key, C is ciphertext. Shannon wrote

> ``Two methods (other than recourse to ideal systems) suggest themselves for frustrating a statistical analysis. These we may call the methods of *diffusion* and *confusion* In the method of *diffusion* the statistical structure of M which leads to its redundancy is dissipated into long range statistics—i.e., into statistical structure involving long combinations of letters in the cryptogram. The effect here is that the enemy must intercept a tremendous amount of

[20] Originally issued as a classified paper on September 1, 1945.

[21] It was known that *one-time keys* achieve perfect secrecy; it was invented in 1882 and reinvented in 1917 by Gilbert Vernam of ATT. Perfect secrecy requires that the number of messages is finite and to be the same as the number of possible keys. If the message is thought of as being constantly generated at a given, the key must be generated at the same or a greater rate.

material to tie down this structure, since the structure is evident only in blocks of very small individual probability."

- Shannon offered the (keyless) auto-encipherment example of diffusion

$$\underline{m} = (m_0, m_1, \ldots, m_{n-1})$$

$$\underline{c} = (c_0, c_1, \ldots, c_{n-1}) \qquad c_j = (m_j + m_{j+1} + \ldots + m_{j+n-1}) \text{ (mod 26)}$$

- The Hill Cipher[22] is also cited by Shannon as an example of diffusion.
- A *transposition (*or *wire crossing*), namely a permutation of the bits of a plaintext sequence, is an example of diffusion.

  ``The method of *confusion* is to make the relation between the simple statistics of C and the simple description of K a very complex and involved one."

Section 25 of Shannon's secrecy paper combines *confusion* and *diffusion with* a *mixing transformation* as defined in ergodic theory [4]. Shannon writes

> ``Speaking loosely, however, we can think of a mixing transformation as one which distributes any reasonably cohesive region in the space fairly uniformly over the entire space. Good mixing transformations are often formed by repeated products of two simple non-commuting operations."

3.4 HORST FEISTEL'S WORK AT IBM[23]

1. Horst Feistel is the author of a very well cited paper [22] on the role that cryptography could play in securing privacy in computer systems.
2. IBM filed several patents[24] based on his cryptographic ideas including

[22] In traditional cryptography, the Hill cipher is a polyalphabetic substitution, based on the linear equation C = K M (the key K is an n-by-n matrix; M is an n-vector). Proposed by Lester S. Hill in 1929, it was the first practical polyalphabetic cipher to encipher more than three symbols at once.

[23] Photograph circa1971 courtesy of the IBM Corporation

- Block Cipher Cryptographic System, US#3798359A, filed on June 30, 1971;
- Key Controlled Block-Cipher Cryptographic System Employing a Multidirectional Shift Matrix, US# 4195200A, filed on June 30, 1976;
- Stream/Block Cipher Cryptographic System, US#4316055A, filed on December 30, 1976.

3. While `FORTRAN` was created in 1954 at IBM, Horst went his own way, writing his encryption programs in APL[25] (*A Programming Language*). Horst intended to name the program implementing his encipherment algorithm DEMONSTRATION, but early versions of APL limited the character length of a file name and `DEMON` was the natural replacement. His colleague Lynn Smith may have suggested `LUCIFER` and Horst concurred, probably sensing that it was a sexier choice.

3.5 IBM ENTERS THE CRYPTOGRAPHY BUSINESS

The IBM Corporation has long sponsored and encouraged the patenting of its research as stated on its corporate webpage.

> ``IBM has focused on continuous innovation for more than a century. Patenting is an important barometer of that innovation, and IBM has topped the annual list of U.S. patent recipients for the 20th consecutive

[24] In addition, the US #3,962,539 ``Product Block Cipher System for Data Security" describing the design of DES was filed by IBM Kingston on June 28, 1976. Listed as inventors were William Friedrich Ehrsam, Carl H. W. Meyer, Robert Lowell Powers, John Lynn Smith and Walter Leonard Tuchman.

[25] Kenneth E. Iverson developed APL in the 1960s. It was marketed as a Program Product by IBM and had an important influence on the development of spreadsheets, functional programming and computer- oriented mathematical programming packages (MatLab, Mathematica, and Scratchpad)...

The free-software activist Richard Stallman is the author of this ditty, to be sung to the melody of ``Row, row, row your boat gently down the stream.`

Rho, rho, rho of X Always equals 1
Rho is dimension, rho rho rank.
APL is fun!

I am a great fan of APL and perfected my text editing skills while at IBM by photocomposing Allen J. Rose's book on *APL*. I also taught Computer Science 5 (APL) twice at the UCSB Micro Lab on Apple computers.

APL may be infrequently in use today, but *great* ideas always survive; to wit, GNU offers 32- and 64-bit implementation for Windows.

> year. From 1993-2012, IBM inventors received nearly 67,000 U.S. patents, and in 2012 alone, received a record 6,478 patents, exceeding the combined totals of Accenture, Amazon, Apple, EMC, HP, Intel, Oracle/SUN and Symantec.
>
> Most important, these inventions have a huge impact on clients, partners and society. They show IBM's long-term, strategic commitment to innovation and demonstrate the patience to allow scientific discovery to find its way into the market."

Horst Feistel's research might just have been just *blue sky* and not found any commercial application, except that IBM now entered the crypto-business. IBM began the development of their 2984 (Cash Issuing Terminal) in 1966 because of their contract with the *Lloyds Banking Group*. The IBM 2984 became an early component of the *Lloyds Bank Cashpoint System*. Now referred to as an ATM (*Automated Teller Machine*), an IBM 2984 became operational in 1972 in Essex England.[26]

The Systems Communication Division (SCD), located along the Hudson River in Kingston (New York), was assigned responsibility for the crypto-product development. An ATM user needs to be authenticated before any transaction can take place. Worried about the risks of dispensing cash like chewing gum or cigarettes, the banking community insisted that a user be authenticated prior` to an ATM transaction by using two tokens; by validating the relationship between the first token, the *primary account number* (PAN) written on the user's banking card, and the second token, the personal identification number (PIN), entered by the user at a keyboard. Cryptography

[26] IBM WHITE PLAINS, N. Y. June 12, 1972

International Business Machines Corporation announced a computer terminal that reads credit cards and issues $5, $10 or $20 bills today.

The IBM 2984 cash issue terminal, which operates under the control of a bank's central computer, permits the holder of a valid credit card to withdraw cash night or day, and without having to write a check. It can be installed on the outside of a bank for 24-hour use or in the lobby to speed withdrawals during busy periods.

Inserting a credit card opens a panel covering two keyboards. A personal identification number is entered on one keyboard and the amount of cash desired on the other. This information, along with account data encoded invisibly on the card's magnetic stripe, is transmitted to an IBM System/370 or System/360, which checks the validity of the transaction. If valid, the cash is dispensed is used to validate the PIN/PAN. As a first step, IBM SCD needed to identify an appropriate encipherment algorithm.

Like Yorktown Research, the Kingston group was just acquiring technical competence in cryptography. SCD initially considered the Hill cipher to provide the connection between the PIN and PAN in ATM transaction. While Hill encryption has a potentially large key space, its encryption is a linear transformation and therefore susceptible to a (partial) known *plaintext* attack. Walter Tuchman, the IBM project manager quickly realized the weakness of the Hill Cipher. As Hesiod (~800 BC), the Greek didactic poet recognized, ``timing is everything;" LUCIFER was available and IBM and SCD decided to modify it, referring internally to it as DSD-1. Close technical cooperation developed between the Yorktown and Kingston groups in the process. DSD-1 incorporated in the IBM 2984[27] became DES, a Federal information Processing Standard (FIPS) in 1976. The IBM Kingston group included Carl H. Meyer and Stephen M. Matyas whose book [39] describes some aspects of the DES design.

Having decided to enter the cryptographic design business, the IBM Corporation wanted to avoid the embarrassment which would result if their cryptographic algorithm was shown too weak and became the IBM Corporation's Edsel. To try to avoid this pitfall, IBM Yorktown Research engaged the services of two consultants with *prior* experience in cryptography. The career of Professor Edward L. Glazer from Case Western Reserve University involved computer security research. He was a member of the National Academy of Engineering and the National Science Foundation. The second consultant was Professor James Simons[28] from SUNY Stony Brook. Jim Simons had` previously worked at the Communications Research Division of the Institute for Defense Analysis (IDA/CRD). Located in Princeton (New Jersey), IDA/CRD has been an NSA contractor since its founding in 1959[29]. IDA/CRD sought to enlist the talents of

[27] DES hardware description: ~ ¼ by ¼ inch, 2 micron CMOS, containing one DES-engine and enciphering at the rate of 4 Mb/s (64 bits in 16 microseconds).

[28] In 1982, Simons founded Renaissance Technologies, a private hedge fund investment company based in New York with over $15 billion under management. Simons retired at the end of 2009 as CEO of one of the world's most successful hedge fund companies. Simons' net worth was estimated to be $12.5 billion. It has been suggested that Simons applied the Hidden Markov Model methodology to guide his investment. Perhaps, the study of algebra, probability and statistics might lead to wealth as well as fame!

[29] Nancy Albert’s book cited before mentions that IDA was founded in 1956 and in a sense it is an outgrowth of the Bowdoin College meeting sponsored by Cambridge AFCRC. During 1961-62, A. A. Albert took a leave of absence and served as the director of IDA/CRD.

American mathematicians by sponsoring the SCAMP (Summer Conference on Applied Mathematics Problems); the word *Problems* referred to those related to the cryptography area.

What was the result of our interaction with the consultants? Levy [33, p. 48] cites me in relating Glazer's unsubstantiated claim that he could *break* `LUCIFER` with 20 pairs of corresponding plaintext and ciphertext. He was not successful. James Simons prepared a report for IBM writing in its summary, ``… no attack was found which would recover key from arbitrary amounts of matched plain and cipher. Neither was an attack discovered that would enable one to read a significant fraction of messages without a work factor comparable of trial and error."

Although, it was rumored NSA suggested that IBM embed a *backdoor* in its algorithm to facilitate deciphering its ciphertext, neither Glazer nor Simons found any evidence of this.

Those allowed to peak under the *Cone of Silence*[30] are not encouraged to market their expertise. I discovered that this is also true of non-American experts. My first trip to California occurred in 1961 to attend a Summer Conference on Functional Analysis at Stanford University. I attended a lecture there on functions of several complex variables given by Professor Arne Beurling, a permanent member of the Institute for Advanced Studies in Princeton until his death in 1986. During World War II, Professor Beurling reverse-engineered one version of the Siemens and Halske T52 *Geheimschreiber*, developing a cryptanalysis for it. I invited Professor Beurling to give a lecture at IBM Research, indicating our interest in his cryptanalysis work. Although Beurling had no connection with NSA, his prior relationship to the Swedish NSA, the National Defense Radio Establishment (FRA), muted his voice. If IBM Research wanted to learn whether DES was strong or weak, the Mathematical Sciences Department would have to figure it out on our own.

The IBM Science Advisory Committee met in May 7, 1973[31]. They implemented two policies; *i)* the IBM would have a single cryptographic architecture and *ii)* that a

[30] In my conversation with Ms. Nancy Albert, she stressed how reluctant her father was to talk about *any* aspect of his cryptography work.

[31] IBM Confidential document provided courtesy of the IBM External Submissions Office.

cryptanalytic capability would be needed to assure ``adequate technical contention in cryptography."

3.6 THE CONFLICT WITH NSA

IBM's long standing practice was to protect the technology it developed by filing a U.S. patent application; in this instance, the use of cryptography in the banking automated teller system. Since this application originated from a contract with the Lloyds Banking Group in London, it made sense for IBM, as an international corporation, to file a European patent application.

The Invention Secrecy Act of 1951 (35 USC§§ 181-8) is a body of United States federal law designed to prevent disclosure of new inventions and technologies that, in the opinion of selected federal agencies, present a possible threat to the national security of the United States. Before foreign patent coverage can applied for, a patent application must be first filed in the United States and reviewed by government agencies selected by the Patent Office.[32] The Commissioner of Patents issues a secrecy order to stop the patent process in instances, if the publication of an application or the granting of a patent may be detrimental to national security.[33]If in the six months after the submission of a patent application in the United States no secrecy order results, patent applications outside the United States may proceed.

Since the patent title incorporated the term *cipher,* NSA quickly realized they had an interest in the IBM patent application US#3798359A, filed on June 30, 1971 with Horst Feistel named as the inventor. The conflicting issues at play were

i) denying adversaries of the United States new cryptographic technology, in order to better to protect secure the country, and
ii) the potential business opportunities to IBM offered by current or future products relying on cryptography marketed by them[34]

[32] 35 U.S.C. 184 Filing of Patent Application in a Foreign Country.

[33] These rules, implemented before world-wide email and the Internet, seem nonsensical today,

[34] In a March 27, 2003 article ``How the ATM Business Revolutionized Banking "in Bloomberg News, the author Bernardo Batiz-Lazlo summarizes the history of ATMs in banking. The explosion in banking is due in part to the use of dial-up connections replacing dedicated lines making ATMs more economical.

It took NSA nearly a year and a half, but it resulted in a secrecy order issued on October 17, 1973. I have no recollection of the substance of the rumor that NSA severely *edited* one IBM Yorktown paper as a prerequisite for publication. Since Horst Feistel had let the cat out of the bag, the secrecy order seemed ludicrous. Nevertheless, the government continued to push for secrecy. Victor Siber, an attorney in the Intellectual Properties Office in Yorktown at this time, suggests that it might have been used for some bargaining advantage[35].

November 14, 1973 witnessed the lifting of the secrecy order. While IBM had been free to file a European patent application on any time between January 30, 1972 and October 16, 1973, a patent search on Google does not reveal any filed applications during that period. Victor Siber states that IBM filed several European applications, but explained the patent review process is lengthy and IBM decided to withdrew the application when the secrecy order issued.

Even though very clever people manage NSA, it took many years before they realized the silliness and futility of the key-size limitation. While Europe is only a *pond* away from the United States, most places on earth are only a few milliseconds away. Before DES, the Export Control Act limited key length to 40 bits; DES used 56 key bits. In January 2000, all key-length restrictions disappeared for DES and its successors. Hardware devices containing DES were exportable to all except countries other than members of President Ronald Reagan's 1983 *axis of evil*. In the interim, there might have been real or imagined attempts to *censor* the academic community, whatever those terms mean. Richard L. Garwin[36] is a distinguished American physicist and an IBM Fellow Emeritus. He counseled IBM not to force pre-review before allowing publication and they took his advice.

[35] Walt Tuchman stated in email that the key size limitation resulted from the chip design limitations. He believed the most important issue for NSA was the secrecy of design criteria; in particular, the T-attack found by IBM Research. Twenty years later, it was described in the Don Coppersmith's paper [27].

Eli Bilham (Differential Cryptanalysis) and a related attack by Mitsuru Matsui rediscovered the T-attack in 1990 in 1993 (*Linear Cryptanalysis*).

[36] *Wikipedia* claims that Richard L. Garwin was connected with the design of *Mike,* the first hydrogen bomb in 1952. 💣

Victor Siber suggested in a conversation that it was in IBM's interest to have DES new Federal Standard for digital encryption. Moreover, IBM issued a public notice (in the USPTO Gazette) that the public was given a royalty free license under three U.S. cryptography patents , including Feistel's 1971 patent, for making, using and selling DES in the U.S

ATMs are now ubiquitous throughout the world and represent the *first* real successful commercial application of encipherment. Yahoo Finance reported in April 2014 that China for the first time exceeded the United States as the country with the largest number of ATMs.

The IBM-NSA dispute originated with the use of cryptography in ATM transactions in the banking industry. E-Commerce[38] is a potentially more profitable commercial application of encryption than banking is. Cryptography functions in an ATM to provide the *authentication* of a user at a secured location (a bank or retail location). E-Commerce has to support *authentication* of one or both parties linked over the insecure Internet. In 1970, the tools for E-Commerce remained to be developed; some today believe even more work is required.

3.8 WHAT HORST FEISTEL DID NOT DO

1. Horst did **not** invent *public key cryptography* (PKC).

The concept of PKC first appeared in the unclassified literature in the celebrated1976 paper [16] authored by Whitfield Diffie and Martin Hellman. The public key or two-key concept appeared earlier in the *classified*[39]papers of Cocks [13] and Ellis [20], who worked for the Government Communications Headquarters (GCHQ).[40] Neither Cocks nor Ellis realized the significance of PKC and its potential commercial application. On the other hand, this was the motivation and starting point for Diffie and Hellman.

[38] E-Commerce (*noun*): Commercial transactions conducted electronically on the Internet.

[39] The original papers were declassified in 1997.

[40] GCHQ is the British analog of NSA the intelligence and security agency responsible for providing signals intelligence (SIGINT) and information assurance to the British government and their armed forces.

The Diffie-Hellman paper however did not provide a viable example of a public key system. Professor Donald Knuth had suggested to Diffie, who was then a graduate student at Stanford, that factorization is an example of a one-way function. All the same, implementing a PKC from a one-way is far from obvious; for whatever reasons, Diffie and Hellman chose not to further explore this direction. More significantly, the Diffie-Hellman paper also conjectured that PKCs could permit the construction of *digital signatures* to demonstrate the authenticity of a message or document. This concept has found great applicability in Internet-based commercial agreements; for example, in the signing of contracts between remotely connected parties.

The knapsack problem is also a one-way function; Merkle and Hellman [38] constructed a very ingenious PKC based on it. Adi Shamir [52] provided a remarkable and very elementary cryptanalysis; it was announced first in at the 1982 Crypto '82[41] conference at UCSB. Adelman [1], Lagarias [32], and Lagarias and Odlyko [33] published extensive connections between knapsack-based cryptosystems and Diophantine approximation.

The discovery of a true PKC had to wait until the appearance of the RSA paper [51]. Victor Miller's paper [40] and Neal Koblitz [30] both explained how the *chord-tangent group law*[42] provided a PKC. Like RSA, it was based on the one-way additive version of *factorization* problem situated on elliptic curve groups over finite fields. An explicit elliptic curve group PKC is described by Menezes and Vanstone [37].

2. Horst did **not** investigate the *ATM PIN/PAN Protocol* or contribute to the *ATM Bankcard Standard.*

Geoffrey Ernest Patrick Constable [14] is the inventor named on US 3,543,904, a patent issued in 1970 to Smith Industries Limited. Constable describes the basic ATM PIN/PAN protocol. An ATM incorporating it was operational at the Westminster Bank in England by Chubb Integrated Systems in May 1968. The Constable protocol has long

[41] UCSB continues today to be the site of a meeting in August focusing on current topics in cryptography. CRYPTO 2014 was held at UCSB from August 17-24, 2014.

[42] Discovered in the nineteenth century by Carl Gustav Jacobi.

since been replaced by Martin Attalla's protocol[43].

The ATM authentication protocol uses data (PIN) entered by a customer at a keyboard and data (PAN) read from the third track of the banking card, the format specified in the ANSI standard [3]. Various attacks on the PIN/PAN protocol have been evaluated [9], excluding the traditional (decipher-proof) armed robbery of a customer at an ATM terminal machine site.

3. Horst Feistel certainly understood some aspects of authentication as evidenced in his 1958 MIT Lincoln Laboratory Report [21]. Nevertheless, he did **not** offer a solution of these principles as they apply in the Internet today to the remote authentication of a user as described in (X.509) [15].
4. Horst did **not** contribute to the important enabling mechanism for today's E-Commerce; for example, the *Secure Socket Layer Protocol* (TLS) [19].

Authentication of a bank customer **at** a normally secure ATM terminal is the basic requirement in an ATM transaction. The authentication on the Internet of a vendor to a remotely connected customer is the necessary requisite in E-Commerce transaction. Items 1, 3 and 4 are prerequisites for secure E-Commerce. Both secrecy and authentication, as practiced today in E-Commerce, require a strong encryption algorithm.

3.9 WHAT HORST FEISTEL DID ACHIEVE?

1. He invented the first of a series of 20th century cipher algorithms [23, 57 and 58].
2. DES implemented at SCD Kingston was derived from the ideas contained in `LUCIFER.` NBS[44] published requests in the Federal Register for an encryption algorithm in 1973 and again in 1974. IBM responded, the algorithm was published and comments solicited in 1975. Two workshops in

[43] Born in Egypt, Atalla earned the M.S. (1947) and PhD (1949) degrees in mechanical engineering at Purdue University and worked at Bell Labs. After leaving Bell Labs, Atalla co-founded Hewlett-Packard Associates to provide them with solid-state capabilities. In 1973, he founded his own company, the Atalla Technovations Corporation, to address the security requirements of banking and financial institutions. He invented the so-called "Atalla Box", a security system that secures a majority of transactions from ATMs today.

[44] Founded in 1901, the *National Bureau of Standards* (NBS) was a repackage as the *National Institute of Standards and Technology* (NIST) in 1988.

Gaithersburg, Maryland were held in 1976 to evaluate the technology and mathematical foundation of DES. DES was first approved in November 1976 and published as the Federal Information Processing Standard (FIPS) in January 1977 [41] and reaffirmed a standard several times [42, 43], most recently in 1999 [44].

The January 2, 1997 issue of the Federal Register contained the statement

> At the next review (1998), the algorithm specified in this standard will be over twenty years old. NIST will consider alternatives, which offer a higher level of security. One of these alternatives may be proposed as a replacement standard at the 1998 review.

3. *Rijndael* [17] was the successor to DES and the winning algorithm in October 2000; it was affirmed as the new standard, the *Advanced Encryption Standard* (AES) [45].

4. These two new cryptographic algorithms differed from the 19th and early 20 th century cipher machines. Until the 1971, encryption used either shift-register stream encipherment natural for voice or electro-mechanical cipher machines; an excellent description of the machine devices is given in the book by Deavours and Kruh [18]. The use of electro-mechanical machines[45]resulted perhaps in some algebraic structure in the ciphertext, for example, the cycle structure. It was a starting point in their mathematical cryptanalysis. In addition, the *Enigma* ciphertext enjoyed the property that no plaintext letter could be enciphered to itself. While this does not seem much, it was.

5. Horst Feistel's block cipher patent provided the impetus for investigations that have inexorably led technology to today's E-Commerce!

6. The two major missions of NSA are *Information Assurance* (IA) and *Signals Intelligence* (SIGINT) ; IA protects our country's communications, SIGINT is described thusly on their website, writing that NSA

   `` … collects, processes, and disseminates intelligence information from foreign signals for intelligence and counterintelligence purposes and to support military operations.`

[45] Rotors, pin wheels, lugs, cipher wheel, disks, and rotary dial telephone switches.

Horst Feistel's work has unintentionally and *vastly* complicated today the NSA SIGINT mission; `LUCIFER`, DES and AES were the first of a class of Shannon-inspired encipherment systems. Edgar Allan Poe's story *The Gold Bug* contains the oft-quoted statement

> ``Yet it may be roundly asserted that human ingenuity cannot
> concoct a cipher that human ingenuity cannot solve`

NSA and GCHQ have had an illustrious history involving the successful cryptanalysis of German and Japanese cipher systems during World War II. However, with the introduction of DES and sequels, Poe's statement might not remain accurate today. It might be now too costly for NSA to decipher messages encrypted with DES, AES or their subsequent commercial or foreign intelligence agency successors.

Continued success in their SIGINT mission might depend on NSA employing trapdoor or backdoor attacks.[46] Various websites on the Internet describe many backdoor/trapdoor attacks on cryptographic algorithms and devices that use them; I point to two celebrated examples and note an abundance of other examples.

- The illegal entry in 1941 by the FBI agents and the U.S. Naval Intelligence (OP-20-GY) into the Japanese consulate in Manhattan to photograph codebooks [60].
- The agreement in 1958 negotiated by the retired American cryptographer William Friedman and Boris Hagelin, the founder and CEO of the Swiss company *Crypto AG.* It secretly allowed NSA to design and insert backdoors in the cryptographic equipment supplied at least

[46] In a legitimate theatre, a trapdoor is a sliding or hinged door, flush with the surface of a floor, roof, or ceiling, or in the stage of a theatre. People seem to appear or disappear in a puff of smoke using the trapdoor, which hides the closing or opening of the hidden door. A cryptographic *trapdoor*, it is an alteration of the enciphering program, which allows the trapdoor inserter's agents to read messages without the sender's knowledge. A cryptologic *backdoor* is a synonym for *trapdoor,* also including a variant involving surreptitious entry into a subject's premises to steal some information

through 1992 to some of Hagelin's customers, intelligence agencies of governments whose policies were decidedly hostile to the United States [53, part 4, pp. 9-11].

- A Google search for ``NSA backdoors" returns a good number of hits.

3.10 COMPARING SHANNON AND FEISTEL

Acknowledged today as a giant in technology, Shannon made several important technical contributions. In addition to the two papers already cited, his MIT master's thesis [54] applied Boolean algebra in the analysis of switching circuits. It was a cornerstone for the design of computers and influenced the much earlier enhancement of telephone call routing systems All modern communications, including those on the Internet, are made feasible (practical and economical) as a result of research spawned by Shannon's *Theory of Communications* paper. While Shannon did not verbalize the concept of *authentication,* he understood the closely related need for secrecy in communications.

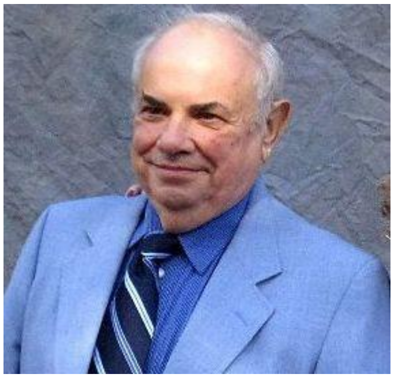
While Horst Feistel was very inventive, he is certainly not in the same technical league as Claude Shannon. Nevertheless, ATM technology and **all** of the advantages of E-commerce are the consequences of the groundbreaking work initiated by Horst Feistel. It is possibly not Feistel's technical achievement that is comparable to that of Shannon, but the impetus it provided for technologies of ATM and E-Commerce that may be.

Perhaps, like all discoveries in cryptography and communications technology, if Shannon or Feistel had not published them in 1948 and 1971, someone else would have ultimately invented it. Conceivably so, but they did.

State Street is the main shopping venue in downtown Santa Barbara, California. Stores open, flourish and sometimes disappear. Already some types of stores have vanished; remaining are banks, national-chain stores, bars, restaurants and coffee shops. *Amazon*.com is not the only E-Commerce winner; even big- and small-box stores market on-line to survive. It will be years before we will be able to decide if E-Commerce is a *blessing* rather than a *curse*. While the social utility of Shannon's work is not an issue, the same judgment might not be true for E-Commerce. It is perhaps too early to tell and I am not percipient enough to give an answer. If E-commerce turns out to be not only a winner for not only the cable providers, USPS, UPS and FedEx, but for society, then … we have (at least) a tie!

I gratefully acknowledge the comments, (constructive) criticism, advice and assistance I have received from many friends, former colleagues and new acquaintances, including

- Ms. Peggy Gertrude Chester (née Feistel)
- Dr. Whitfield Diffie,
- Dr. Richard Garwin (IBM Fellow Emeritus),
- Ms. Susan Greco (IBM External Submissions Office),
- Professor Emeritus Martin Hellman,
- David Kahn, celebrated author and friend,
- Dr. Carl Meyer (IBM Kingston *Emeritus*),
- Victor Siber *Esq.(*retired from IBM)[47],
- Ms. Dawn Stafford (IBM Corporate Archives Reference Desk), and
- Dr. Walter Tuchman (IBM Kingston *Emeritus*).



__Author Information__: Alan G. Konheim attended the Polytechnic Institute of Brooklyn, receiving a B.E.E. in 1955 and a M.S. (Mathematics) in 1957. After completing the PhD program in Mathematics at Cornell University in 1960, he joined the Mathematical Sciences Department at the IBM Yorktown Research Center. In 1982, seeking a sunnier

[47] Victor Siber worked at IBM Research in the IP capacity; subsequently, he was Chief Intellectual Property Counsel for IBM. Today, Mr. Siber is a Patent Attorney at siberlaw.com in Manhattan;

climate, he left IBM joining the faculty of the Computer Science Department at the University of California in Santa Barbara. As Nelly Furtado's song explains, ``all good things come to an end'' and Alan Konheim became Professor Emeritus in 2005.

References

1. Adelman, Leonard 1983. On Breaking Generalized Knapsack Public Key Cryptosystems, *Proceedings of the 15th ACM Symposium on The Theory of Computing,* 402–412.
2. Albert, Nancy E. 2005. A^3 His Algebra: How a Boy from Chicago's West Side Became a Force in American Mathematics, New York: iUniverse.
3. American National Standards Institute. *The ANSI Standard X9.1, Bank Cards – Magnetic Stripe Content for Track 3.* 1980.
4. Arnold, I. and Avez, A. 1968. Ergodic Problems of Classical Mechanics, New York: W.A. Benjamin.
5. Bahl, L, Cocke, J., Jelinek, F. and Raviv, J. March 1974. Optimum Decoding of Linear Codes for Minimizing Symbol Error Rate, *IEEE Transactions on Information Theory*, **IT-20**: 284-287.
6. Baum, L.E., Petrie, T., Soules, G. and Weiss, N. 1969. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *Annals of Mathematical Statistics*, **41**: 165–171.
7. Berlekamp, Elwyn 1968. Algebraic Coding Theory, New York: McGraw-Hill.
8. Berrou, C., Glavieux, A. and Thitimajshima, P. May 1993. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes, *Proceedings 1993 International Conference on Communications*, Geneva, 1064–1070.
9. Bond, M. and Zelinski, P. February 2003. Decimalization Table Attacks for PIN Cracking, University of Cambridge, Computer Laboratory, and Report 560.
10. Bose, R. C. and Ray-Chaudhuri, D. K. March 1960. On a Class of Error-Correcting Binary Group Codes, *Information Control,* **3:** 68–79.
11. Christensen, Chris 1998. Review of A_3 and His Algebra by Nancy E. Albert, *Cryptologia*, **32**(12): 189-196.
12. Coppersmith, Donald May 1994. The Data Encryption Standard (DES) and its Strength against Attacks, *IBM Journal of Research and Development,* **38**(3): 243-250.
13. Cocks, C. C. November 20, 1973. A Note on Non-Secret Encryption, CESG Report, GCHQ.

14. Constable, G. E. P. (Filed) March 5, 1968. U.S. Patent No. 3,543,904Access-Control Equipment and Item-Dispensing Systems Including Such Equipment.
15. Consultation Committee, International Telephone and Telegraph (CCITT), Geneva (Switzerland): *Recommendation X.509, the Directory – Authentication Framework* International Telecommunication Union (ITU), 1989.
16. Diffie, Whitfield and Hellman, Martin 1976. New Directions in Cryptography, *IEEE Transactions on Information Theory,* **IT-22**: 644–654.
17. Daemen, Joan March 1995. Cipher and Hash Function Design Strategies Based on Linear and Differential Cryptanalysis, Doctoral Dissertation at K. U. Leuven.
18. Deavours, Cipher A. and Kruh, Louis 1985. Machine Cryptography and Modern Cryptanalysis, Norwood, Massachusetts: Artech House.
19. Dierks, T. and Rescorla, E. August 2008. The Transport Layer Security (TLS) Protocol, Version 1. Network Working Group, RFTM.
20. Ellis, J. H. January 1970. The Possibility of Secure Non-Secret Digital Encryption, CESG Report, GCHQ.
21. Feistel, Horst May 20, 1958. A Survey of Problems in Authenticated Communication and Control, MIT Lincoln Laboratory, 1-111,
22. Feistel, Horst May 1973. Cryptography and Computer Privacy, *Scientific American,* **228**(5): 15-23.
23. Feistel, Horst (Filed) July 30, 1971. US Patent 3,798,359, Block Cipher Cryptographic System.
24. Feistel, H., Notz, W.A. and Smith, J.L. November 1975. Some Cryptographic Techniques for Machine-to-Machine Data Communications, *Proceedings of the IEEE,* **63**(11): 1545-1554.
25. Gorenstein, Daniel 1988. The Classification of Finite Simple Groups, A Personal Journey: The Early Years, Part I: History of Mathematics, *American Mathematical Society,* **1***:* 447-476.
26. Hagenauer, Joachim and Hoeher, P. November 1989. A Viterbi Algorithm with Soft-Decision Outputs and its Applications, *Proceedings IEEE Global Telecommunications Conference* 1989, Dallas, Texas: 1680–1686.
27. Hamming, Richard W. April 1950. Error Detecting and Error Correcting Codes, *Bell System Technical Journal,* (**29**)2: 147-160.
28. Hocquenghem, A. 1959. Codes Correcteurs d'Erreurs, *Chiffres,* **2**: 147–156.

29. Kahn, David 1967. The Codebreakers (The Story of Secret Writing), New York: The MacMillan Company.

30. Koblitz, Neal 1987. Elliptic Curve Cryptosystems, *Mathematics of Computation,* **48**(177): 203–209.

31. Lagarias, Jeffrey 1984. Knapsack Public Key Cryptosystems and Diophantine Approximation, Advances in Cryptography, Plenum Publishing Company New York.

32. Lagarias, Jeffrey and Odlyzko, Andrew. 1985. Solving Low-Density Subset-Sum Problems, Journal *of the Association for Computing Machinery*, **32:** 229–246.

33. Levy, Steven 2001. Crypto: How the Code Rebels Beat the Government – Saving Privacy in the Digital Age, New York: Viking.

34. Lin, Shu 1970. An Introduction to Error-Correcting Codes, Upper Saddle River, New Jersey: Prentice-Hall.

35. Lucky, Robert W.1991.Coding is Dead, *IEEE Spectrum,* 243-246.

36. MacWilliams, F. J. and Sloane, N. J. A. 1977. The Theory of Error-Correcting Codes. Amsterdam (The Netherlands): Elsevier.

37. Menezes, A.J. and Vanstone, Scott A. 1993. Elliptic Curve Cryptosystems and Their Implementation, *Journal of Cryptology***, 6:** 209–224 (1993).

38. Merkle, Ralph and Hellman, Martin 1978. Hiding Information and Signatures in Trapdoor Knapsacks, *IEEE Transactions on Information Theory,* **IT-24**: 525–530.

39. Meyer, Carl H. and Matyas, S. M. 1982. Cryptography: A New Dimension of Data Security, New York: John Wiley.

40. Miller, Victor 1985. Use of Elliptic Curves in Cryptography, *Proceedings of CRYPTO '85,* Springer-Verlag: Berlin, Germany, 417-426.

41. National Bureau of Standards (NBS). *Data Encryption Standard (DES).* FIPS Publication 46. Government Printing: Office: Gaithersburg, Maryland. January 15, 1977.

42. National Bureau of Standards (NBS). *Data Encryption Standard (DES).* FIPS Publication 46-1. Government Printing: Office: Gaithersburg, Maryland. January 22, 1988.

43. National Bureau of Standards (NBS). *Data Encryption Standard (DES).* FIPS Publication 46-2. Government Printing: Office: Gaithersburg, Maryland. December 30, 1993.

44. National Bureau of Standards (NBS). *Data Encryption Standard (DES).* FIPS Publication 46-3. Government Printing: Office: Gaithersburg, Maryland. October 25, 1999.

45. National Bureau of Standards (NBS), *Advanced Encryption Standard (AES)* FIPS Publication 197. Government Printing: Office: Gaithersburg, Maryland. November 26, 2001.

46. Peterson, Wesley W. 1961. Error-Correcting Codes. Cambridge, Massachusetts: MIT Press.

47. Peterson, Wesley W. and Weldon Jr., E. J. 1972. Error-Correcting Codes. Cambridge, Massachusetts: MIT Press.

48. Rabiner, L. R. and Juang, B.H. January 1986. An Introduction to Hidden Markov Models, *IEEE ASSP Magazine,* **3**:4–16.

49. Reed, Irving S. September 1954. A Class of Multiple Error Correcting Codes and the Decoding Scheme, *IEEE Transactions on Information Theory,* **IT-4**: 38–49.

50. Reed, Irving S. and Solomon, Gustave June 1960. Polynomial Codes over Certain Finite Fields, *Journal of SIAM,* **8**: 300–304.

51. Rivets, R., Shamir, A. and Adelman, L. 1978. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, *Communications of the ACM,* **21**(2): 120–126.

52. Shamir, Adi 1984. A Polynomial Time Algorithm for Breaking the Basic Merkle–Hellman Cryptosystem, *IEEE Transactions on Information Theory*, **IT-30:** 699–704.

53. Shane, Scott and Bowman, Tom December 3-15, 1995. No Such Agency (in six parts), *The Baltimore Sun,* 1-16.

54. Shannon, Claude E. 1938. A Symbolic Analysis of Relay and Switching Circuits *Transactions AIEE*, **57**(12): 713–723.

55. Shannon, Claude E. July 1948. A Mathematical Theory of Communication, *Bell System Technical Journal*, (**27**)3: 379-423.

56. Shannon, Claude E. October 1949. The Theory of Secrecy Systems, *Bell System Technical Journal,* (**28**)4: 656-715.

57. Smith, John Lynn (Filed) November 1971. US Patent 3,796,830. Recirculating Block Cipher Cryptographic System.

58. Sorkin, Arthur January 1984. ``LUCIFER: A Cryptographic Algorithm, *Cryptologia*, **8**(1): 22–41.

59. Viterbi, Andrew April 1967. Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm, *IEEE Transactions on Information Theory*, **IT-**13(4): 260–269.
60. Wilford, Timothy January 2002. Decoding Pearl Harbor: USN Cryptanalysis and the Challenge of JN-25B in 1941, *The Northern Mariner/Le Marin du nord*, **XII (**1), 17 - 37.
61. Wolfowitz, Jack 1957. The Coding of Messages Subject to Chance Errors, Illinois *Journal of Mathematics,* 1: 591–606.