논문 2004-41CI-5-4

Hybrid 技法을 적용한 효율적인 신용카드판단시스템

(Anti-Fraud System for Credit Card By Using Hybrid Technique)

조 문 배*, 박 길 흠*

(Moon-Bai Jo and Kil-Houm Park)

요 약

인터넷상의 전자상거래 주문에서 발생하는 수백만건의 트랜잭션 레코드들에 대해 Hybrid 기법으로 데이터마이닝 기술인 연관규칙 탐사기법과 AFS (Anti Fraud System) 를 활용하여 전자상거래 과정에서 흔히 일어날 수 있는 부정 거래를 최소화할 수 있는 새로운 전자결제 신용카드 사기방어시스템을 제안한다. 고객이 웹 상의 거래 콤포넌트에 의한 보안 메세징 프로토콜을 사용하여 거래를 시도하면 과거 트랜잭션 데이터를 이용하여 미리 생성해 둔 사기성 거래에 대한 연관규칙의 적용으로 거래의 위험도를 판단하여 위험도가 높다고 판단될 경우 부가적 신용 정보를 요구하거나 거래를 중단하는 시스템이다. 본 시스템의 장점은 기존의 사기방지시스템 보다 빠른 응답성과 그에 따른 효율성을 들 수 있다.

Abstract

An anti-fraud system that utilizes association rules of fraud as well as AFS (Anti Fraud System) for credit card payments in e-commerce is proposed. The association rules are found by applying the data mining algorithm to millions of transaction records that have been generated as a result of orders on goods through the Internet. When a customer begins to process an order by using transaction components of a secure messaging protocol, the degree of risk for the transaction is assessed by using the found rules. More credit information will be requested or the transaction is rejected if it is interpreted as risky.

Keywords: Anti-fraud Screening System, Address Verification System, Datamining, Association Rule, Direct Hashing and Pruning

I. 서 론

인터넷을 통한 전자 상거래는 익명성의 특성으로 인하여 해킹의 가능성이 많아 전자 결제 및 지불에 대한 사기방어 시스템 (Anti-fraud Screening System) 개발의 필요성이 날로 증대되어지고 있다. 그러므로 안전한 전자 거래를 하기 위해서는 다각적이며 효율적인 대비책을 마련해야 하나, 현실적으론 국내외적으로 인증 표준화 기법의 일부만 채택되어 있는 실정이다[1,2]. 전자거래상의 신용카드 사기를 방지하기 위한 최근의 방법으로서는 주로 데이터마이닝 기법에 기반을 둔 여러 시스템들이 제안되고 있다[3,4,5,6]. 인터넷을 이용한 쇼핑몰의 경우에는 그 특성상 전 세계를 대상으로 거래가 가능하지만 신용정보의 부족과 인증 표준화 적용의 어려움으로 인하여 해외거래 시엔 자국 내로 영업을 한정하는 것보다 훨씬 높은 사기거래를 발생시킬 수가 있어, 전자상거래 확산의 제약점으로 지적되고 있으며 따라서 보다 가볍고 효율적인 사기방지책의 개발이 절실하다.

현재 국제적으로 통용되는 신용카드 인증을 위한 대표적인 방법으로서는 북미에서 적용되고 있는 주소 인증 시스템(Address Verification System : AVS)[7,8] 으로 비자, 마스터카드 그리고 그 외의 주요 신용카드들에 대하여 사용자의 주소를 확인할 수 있는 시스템이다. 사용자 인증을 위해 카드 사용자가 온라인 구매 시 기입하는 Billing Address(결제주소)의 일부분 데이터를

* 정회원, 경북대학교 전자공학과
(School of Electrical Engineering and Computer Science. Kyungpook National Univ.)
접수일자: 2003년12월3일, 수정완료일: 2004년8월31일

카드 발급 은행에 있는 사용자의 주소와 관련된 데이터 파일과 비교하게 된다. 이러한 기존 AVS의 카드 거래 승인을 위한 Payment Gateway의 기술적 한계점은 결제 처리에 따른 공유된 Database에 데이터마이닝 검증을 거치지 않는데 있다[2,9]. 검증을 거치더라도 추출된 위험점수만을 신뢰하는 것이거나[8], 주소의 앞부분 20자리와 ZIP Code의 첫 번째 세 자리 숫자만을 비교하는 매개변수에만 크게 의존하는 분리된 방법이었다. 따라서 이 방법으로는 카드를 이용하는 고객이 정확한 카드 소지자인지를 판별하는 데는 한계가 있다. 왜냐하면, 카드 번호와 가입자 주소를 모두 해킹한 신용 카드사기는 인터넷의 익명성에 의해서 쉽게 AVS 인증 과정을 통과할 수 있기 때문이다. AVS의 또 다른 한계점으로는 전 세계적으로 주소 기입 체계가 똑같지 않기 때문에 여러 나라의 주소 데이터를 공통된 방식으로 코드화하는 데에 많은 시간과 자원이 소요되어 효과적인 사기 방지 시스템을 구축하기가 어렵다는 것이다.

본 논문과 비교되는 VeriSign의 Payment Services[8]의 eFalcon Fraud Model은 훈련 시간이 많이 걸리는 뉴럴 네트워크의 특성상 데이터의 일부만을 표본 추출하여 사용하거나 훈련 예를 잘게 분할하여 사용할 수밖에 없다. 따라서 표본추출 상의 에러가 데이터마이닝 모델 안에 포함되기 쉽고, 더구나 훈련 예를 올바르게 잘게 잘 분해하려면 전문가의 많은 작업 시간이 소요된다. 이에 따라 본 논문에서는 이들의 모델을 개선하여 사용자의 편리성이 보장되면서 비교적 빠른 시간 안에 정확한 본인 인증 모델 을 구축할 수 있는 새로운 효과적인 사기접속을 방지하는 시스템에의 접근을 모색해 보았다. 앞으로 Ⅱ장과 Ⅲ장에서는 데이터마이닝 기법인 연관규칙을 본 시스템에 맞게 변경하여 적용한 방법을, Ⅳ장에서는 이러한 기법을 적용하여 개발된 하이브리드 방식을 사용한 신용카드 신용판단시스템 (Anti-Fraud System : AFS)을, 그리고 결론을 제시한다.

Ⅱ. 연관 규칙

히스토리 데이터베이스의 거래결과에 데이터마이닝 기법을 적용하여 사기거래의 가능성이 높은 연관규칙을 추출하게 된다. 그림 1은 데이터마이닝 기법을 적용할 히스토리 레코드의 한 예이다. 데이터마이닝을 위해서는 각각의 개별 신용정보를 정확하게 유지하는 것이 중요하므로 다차원연관규칙 (Multidimensional Associa-tion Rules) 발견 알고리즘을 응용한다. 다음은 알고리

Order#: 20010118094923
Buyer: Christoper T Cox
Mailing Address: 2397 Rio Dosa Dr.,Lexiton, Kentucky, United Stares
Destination Country: United States
Product Receiver: Christopher T Cox
Shipping Address:2937 Rio Dosa Dr., Lexiton, Kentucky
ZIP/postal code: 40509
Cardholder name: Christoper T Cox
Card#: 5491000914029***
Expiry Date: 1101
Issuing Bank: Fleet
IP Address: 205.188.19*.**
AFS Decision: AA-Authorized, Approved

그림 1. 히스토리 레코드
Fig. 1. History Record.

즘에 대한 간단한 설명 및 본 시스템에서 데이터마이닝 연관규칙 발견법을 적용하기 위해 필요한 정의이다.

$I = \{i_1, i_2, \ldots, i_m\}$을 히스토리 데이터베이스에 축적된 거래 레코드의 각 항목 (Item) 에 대한 집합이라고 하자. T는 트랜잭션 레코드를 모아 놓은 것으로 각 트랜잭션은 고객이 제공한 정보의 항목집합 (Itemset) $X \subseteq I$로 구성된다. 연관 규칙이란 $Y \subset I$, $Z \subset I$, $Y \cap Z = \Phi$일 때 $Y \Rightarrow Z$와 같은 규칙을 말하며, $Y \Rightarrow Z$의 신뢰도가 C%란 항목집합 Y를 포함하는 트랜잭션 중 C%는 항목집합 Z도 역시 포함함을 나타낸다. 한 항목집합 $X \subset I$에 대해 X의 지지율 (Support Ratio) 이란 전체 트랜잭션 중 X를 포함하는 비율을 말한다. $Y \Rightarrow Z$와 같은 규칙의 신뢰도는 {(Y∪Z)의 지지율}/{Y의 지지율}로 계산될 수 있다. 지지율 대신 해당 항목집합이 전체 트랜잭션 중에 몇 번 나타났나를 나타내는 지지수 (Support Number) 를 사용하기도 한다. 어떤 지지율 이상 자주 나타나는 항목을 빈발항목집합(Frequent Itemset) 이라 한다. 항목집합을 구성하는 항목의 수가 n이면 n-항목집합이라고 말한다. 빈발항목집합은 이미 구한 (n-1)-빈발항목집합을 이용해 후보 n-빈발항목집합을 만들고 이를 데이터베이스에서 확인해 n-빈발항목집합을 구하게 된다.

본 시스템에서는 히스토리 데이터베이스에 기록된

```
/* 1단계: 1-빈발항목집합을 구하고 2-항목집합을 만들어 해쉬테이블 H2에 저장 */
s = 최소지지수;
H2의 버킷을 0으로 초기화 /* H2는 해쉬테이블로 2-항목집합을 저장 */

forall 히스토리 레코드 r ∈ D do /* D는 히스토리 데이터베이스 */
    해쉬트리에 각 항목을 삽입하고 각 항목에 대한 count 값을 증가시킴;

    forall 2-항목집합 of x ∈ r do /* 예) t = [1,2,3]이면 x = {1,2}, {1,3}, {2,3} */
        H2[h2(x)]++; /* 해쉬함수 h2를 사용해 해쉬테이블 H2의 적당한 위치에 x를 */
    endfor          /* 삽입하고 해당 항에 대한 count 값을 증가시킴 */

endfor
F1 = {i| i.count ≥ s } /* i는 해쉬트리의 잎노드, F1은 모든 1-빈발항목집합의 집합 */
/* 2단계:k-빈발항목집합을 구하고 (k+1)-항목집합을 만들어 해쉬테이블 Hk+1에 저장 */

k = 2;
Dk = D;
while ( k ≤ given_limit )
    generate_candidate(Fk-1, Hk, Ck); /* 앞에서 구한 (k-1)-빈발항목집합 Fk-1와 해쉬 */
                                      /* 테이블 Hk를 기초로 후보 k-빈발항목집합 Ck를 구함 */
    Hk+1의 버킷을 0으로 초기화; /* Hk+1는 해쉬 테이블 */
    Dk+1 = Φ; /* 초기화; 크기가 줄어들 새로운 히스토리 테이터베이스 */

    forall 트랜잭션 레코드 r ∈ Dk do
        /* r 속에 포함된 Ck의 후보 k-빈발항목집합에 대한 count 값을 증가시키고, r 속에 */
        /* 포함된 k-항목집합에서 k번 이상 나오지 않는 항목은 버림. r'는 새로 만들어진 */
        /* 히스토리 레코드. 즉, r' ⊆ r */
        count_support( r, Ck, k, r');
        if |r'| > k then
            /* (크기가 줄어든) 새로운 레코드 r'의 (k+1)-항목집합에 대하여 해쉬테이블 */
            /* Hk+1을 갱신함. 단, 이때 (k+1)-항목집합의 모든 (k+1)개의 k-항목집합은 Ck */
            /* 에 속한다는 조건이 만족되어야 함. 또 그러한 조건을 만족하는 r'의 항목만을 */
            /* 뽑아 새 트랜잭션 레코드 t''를 생성함. 즉, t'' ⊆ t' */
            make_hasht(r', Hk, k, Hk+1, r'');
            if |r''| > k then Dk+1 = Dk+1 ∪ {r''};
        endif

    endfor
    Fk = {i ∈ Ck| i.count ≥ s };
    k++; /* 계속 반복 계산 */

endwhile
```

그림 2. 연관규칙 알고리즘
Fig. 2. Association Rule Algorithm.

자료의 특성상 긴 연관규칙보다는 실용적인 면에서 길이가 2 내지 3인 연관규칙을 효율적으로 찾는 것이 중요하다. 따라서 Iteration의 초기에 후보 빈발항목집합의 수가 여타 다른 연관규칙 발견 방법에 비해 대폭적으로 줄어 든 해쉉 기법에 기반을 둔 DHP (Direct Hashing and Pruning)[11] 방식을 본 시스템의 응용분야에 맞게 수정하여 그림 2와 같이 적용한다. 대량의 자료를 대상으로 하므로 원 알고리즘에 비해 3단계가 생략되었고 다차원 방식으로 연관규칙을 구한다는 점이 다르다. 알고리즘 1단계에서는 효율적인 Counting을 위해 해쉬트리(Hash tree)를 사용한다. 각 항목이 해쉬트리에 이미 존재하면 그 항목의 지지수는 1이 증가되며, 존재하지 않으면 그 항목을 지지수 1을 갖게 하면서 해쉬트리에 삽입하게 된다. 위 알고리즘의 2단계에서는 대상 트랙잭션 레코드의 크기 및 수가 줄어들 수 있음에 주의해야한다. 각 단계에서는 보다 빠른 작업처리를 위해 해쉬테이블을 사용한다. 알고리즘에서 나타낸 해쉬테이블 Hk의 첨자 k는 각 버킷의 크기이다. 아울러 각 버킷에는 이제까지 해당 버킷에 해쉉 되어진 항목집합들의 개수를 나타내는 수를 포함한다.

III. 다차원 연관규칙

원 연관규칙 알고리즘이 항목을 구성하는데 있어 레코드 내의 속성 (Attribute) 구분이 필요 없는 일차원적 (Single-Dimensional, Intra-Attribute) 빈발항목집합의 발견법 이라면, 항목은 값 뿐 아니라 속성을 포함하며 같은 속성끼리는 항목집합을 구성할 수 없는 다차원적 (Multi-Dimensional, Inter-Attribute) 연관규칙 발견법을 적용해야 한다[12]. 즉, 항목은 속성-값 (Attribute-Value) 의 쌍이 된다. 단, 주소의 경우 보다 의미 있는 규칙을 발견하기 위해 주소를 구성하는 각 부항목 (Sub-Item)을 값으로 사용한다. 예를 들어, 앞에서 언급한 레코드의 경우 'shipping_address-"2397 Rio Dosa Dr.",'shipping_address-"Lexiton",' 'shipping_address-"Kentucky, United States"'처럼 항목을 만들 수 있다.

DHP 알고리즘에서는 후보항목집합을 만들기 위해 generate_candidate()를 사용하는데 다차원적 연관규칙을 효율적으로 발견하기 위해서는 속성이 서로 다른 항목만이 항목집합을 만들 수 있도록 수정하여야 한다.

다음은 속성을 검사하는 부분이 추가된 generate_candidate() 알고리즘이다. 같은 속성을 갖는 항목끼리는 항목집합을 구성하지 않음을 검사하는 부분이 추가됨으로써 얻는 계산상의 절약은, 만일 각 속성 당 평균 n가지의 값으로 구성된다고 가정하면 아래 알고리즘의 Select 문에서 선택된 항목집합 당 n(n-1) 가지의 항목집합이 덜 생성되는 것이다. 그림 3은 generate_candidate() 알고리즘을 나타낸 것이다.

발견된 항목집합으로부터 규칙을 생성하기 위해서는 다음과 같은 $I = \{i_1, i_2, \ldots, i_m\}$을 히스토리 데이터베

```
generate_candidate(F_{k-1}, H_k, C_k)
  C_k = φ; /* 초기화 */
    forall 빈발항목집합 ∈ F_{k-1} do
    Select 항목집합 c = i_1, i_2, ..., i_{k-1}, j_{k-1}
    from F_{I,(k-1)}, F_{J,(k-1)} where
i_1 = j_1, ..., i_{k-2} = j_{k-2}, i_{k-1} < j_{k-1}, attr(i_{k-1}) ≠ attr(j_{k-1})
      // F_{I,(k-1)}은 F_{k-1}에 속하는 I번째 항목집합,
      // i_{k-1}은 그 항목집합의 (k-1)번째 항목
      // attr(i)는 항목 i 의 속성
      if (H_k[h_k(c)] ≥ s) then C_k = C_k ∪ {c};
  endfor
end generate_candidate
```

그림 3. generate_candidate() 알고리즘
Fig. 3. generate_candidate() Algorithm.

이스에 축적된 거래 레코드의 각 항목 (Item) 에 대한 집합이라고 하자. T는 트랜잭션 레코드를 모아 놓은 것으로 각 트랜잭션은 고객이 제공한 정보의 항목집합 (Itemset) X ⊆ I로 구성된다. 연관 규칙이란 Y ⊂ I, Z ⊂ I, Y ∩ Z = Φ일 때 Y => Z와 같은 규칙을 원칙을 적용한다. X, Y, Z ⊂ I일 때 아래와 같은 두 규칙이 발견되었다고 하자.

$$X, Y \Rightarrow Z \text{ with } C\% \quad (1)$$

$$X \Rightarrow Z \quad \text{with } C'\% \quad (2)$$

만일 C ≤ C'이면 규칙 (1)은 규칙의 왼쪽편의 값 (LHS, Left Hand Side) 가 더 세분화되었음에도 불구하고 규칙 (2)에 비하여 규칙의 신뢰도가 더 작거나 같은 것이 된다. 즉 간단한 규칙이 더 신뢰성이 높고 그 규칙에 의해 커버되는 레코드도 많아 통계적 신뢰성이 높아질 가능성이 있다. 따라서 규칙의 신뢰성이 유사하거나 작으면서 LHS가 다른 규칙의 수퍼 셑 (Super Set) 이 되는 규칙은 제거(Pruning) 되어야 한다. 히스토리 데이터베이스는 예로 든 레코드에서 보듯이 각 필드의 내용이 키 (Key) 적인 성질이 강하다. 또한 마이닝 시스템의 신뢰도 향상을 위해 최소 지지수를 비교적 작은 수 (Number) 로 함에 따라 많은 수의 연관 규칙이 발견된다. 따라서 발견된 연관규칙은 나중의 검색의 신속성을 위하여 규칙의 LHS 및 속성에 따라 인덱스 파일의 형태로 저장해 놓게 된다. 그림 4는 최종적으로 도출되어진 연관규칙의 예이다. 즉, 이와 같은 연관규칙은 Negative Database의 지식베이스로 저장된다. 하나의 거래 트랜잭션에 대해 여러 개의 대응되는 연관규칙이 발견될 때에는 그 중 사기거래의 가능성이 가장 높은 규칙을 적용하게 된다. 잠재적 거래에 대해 데이터마이닝 시스템에서 검사 결과, 사기율이 일정 %를 넘을

```
        . . .
        . . .
Buyer-"Yuriy Baranovskiy" => fraud, 93%
Buyer-"Gianluca Omodeo" => fraud, 44%
        . . .
        . . .
Shipping_address-"Pacific View Flat 10B Tower2" => fraud,
100%
Shipping_address-"Kiev, Ukraine" => fraud, 62%
        . . .
        . . .
```

그림 4. 도출되어진 연관규칙
Fig. 4. Association Rule Result.

경우 순차적으로 IP주소 추적 시스템, 전화번호추적시스템, 카드발급은행지점추적시스템, 빈도 수 체크 시스템, 위치 체크 시스템 등의 프로그램을 통해 순차적으로 인증 과정을 거친다.

Ⅳ. 하이브리드 기법을 적용한 신용카드 판단시스템 (Anti- Fraud System)

하이브리드 기법을 적용한 신용카드 신용판단시스템 (Anti- Fraud System : AFS) 은 혹시라도 연관규칙 알고리즘이 적용된 데이터마이닝시스템을 뚫고 들어올 수 있는 신용카드 사기 행위를 인증 과정에서 빠르게 방어막을 형성하자는 입장에서 구축된 시스템이다. AVS를 통과하여 승인된 데이터는 주소체계가 일정하지 않는 조건인 북미지역이 아닐 경우엔 그 정확도가 떨어져 AFS에서 다시 처리가 되어야한다. 실시간에서 보다 빠르게 처리할 수 있도록 하기 위해 연관규칙 알고리즘 방식의 데이터마이닝을 통해 얻은 사기율이 일정 % 이상일 때 순차적 (Sequence) 인증 처리로 넘겨져서 검증이 이루어지게 하였다. 이러한 Hybrid 방식은 데이터마이닝 기법과 상호보완적으로 작용하여 빠른 인증 속도와 동시에 정확한 인증을 할 수가 있다.

순차적 인증 처리에서는 단계별로 신용상태를 파할 수 있고 아울러 거래마다 히스토리 데이터베이스에 저장된 정보로 남아 개개인의 인증에 보다 능동적이며 효과적인 사기방지시스템을 설계할 수가 있어 시간이 갈수록 Hybrid System은 정확하고 빠른 인증이 가능하게 된다. 그림 5는 이러한 AFS의 인증 절차의 구조를 나타낸 것이다. 예를 들면, IP주소추적시스템에서는 고정적 IP주소를 쓰고 있는 경우 일정 횟수 이상의 신용구매가 성공적으로 이루어졌으면 IP주소 일치만으로 즉시 인증이 되며, 그렇지 않을 경우 단계적으로 질문을 받

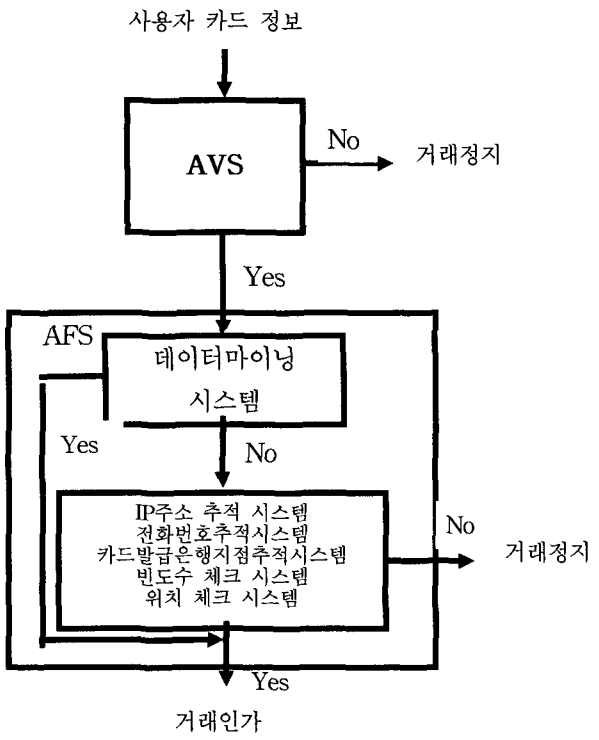

그림 5. AFS의 구조
Fig. 5. Structure of AFS.

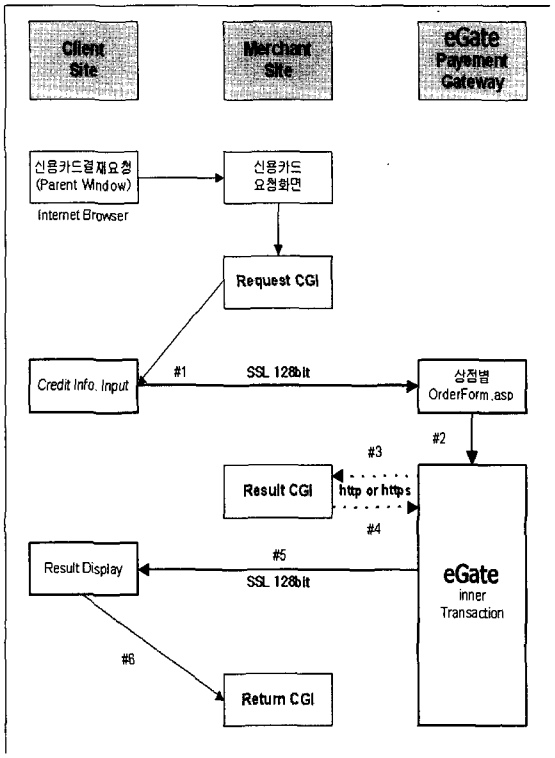

그림 6. AFS 처리를 위한 Data Packet의 진행절차
Fig. 6. Data Packet Process for AFS.

아 거래 히스토리 데이터베이스의 데이터와 비교하게 된다.

고객은 거래 콤포넌트에 의한 보안 메세징 프로토콜을 사용하여 인터넷에서의 거래를 시작한다. 거래의 위험도를 결정하기 위해서 데이터마이닝 기법을 이용하여 이와 같은 요구에서 생성되는 거래자 카드 정보의 위험수위도를 판단한다. 입력된 카드 정보가 사기 거래의 가능성이 높은 것으로 판정될 경우 그 위험도에 따라 IP주소라든가 전화번호 그리고 카드발급은행지점 등을 추적하여 사기거래를 차단하게 된다. 이러한 각 컴포넌트의 기능을 살펴보면

- IP주소 추적시스템은 사용자의 IP주소를 추적하여 본인이 기입한 Billing Address와 Shipping Address 정보가 유용하지 않을 경우 카드 승인을 보류하게 된다. 그러나 선물 처럼 다른 지역에 배달되는 상품일 경우도 있으므로 본인 확인을 위해 전화번호나 카드 발급은행지점 같은 부가적인 신용정보를 요구하게 된다.

- 전화번호 추적 시스템은 사용자가 사전에 입력한 전화번호와 현재 카드 사용자의 전화번호를 대조하여 카드 사기 위험을 사전에 방지해준다.

- 카드발급 은행 추적 시스템에서는 발급 은행 지점 이름을 기입하도록 하여 틀릴 경우 승인을 취소하게 된다.(질문)

- 빈도수 체크 시스템은 카드 사용자가 평소에 카드를 자주 사용하지 않는 사람일 경우 갑자기 쇼핑몰에 접근하여 고액 구매나 많은 양의 주문을 시도한다면 사전에 카드 사용을 막고 카드 사용자 (카드발급 당사자) 에게 한번 더 카드 사용 여부를 파악하는 시스템이다.

- 위치체크시스템은 주로 한 지역에서만 구매를 했던 사람이 동시에 다른 지역에서 구매를 시도할 경우 거리 및 시간을 파악하여 구매를 취소시키는 시스템이다.

AFS는 eGate Inner Transaction Server에서 처리되는데 Merchant Site 와 Client Site로의 연결은 보안 처리된 네트워크를 통해 신용카드 요청화면을 처리하여 연결시켜주는 Request CGI 와 은행망에 연결된 AVS 처리결과를 보내주는 Result CGI, 마지막으로 AFS 처리가 되어 최종적으로 구매자에게 전달되는 Return CGI로 구성된다. 그림 6과 7은 도출되어진 연관규칙을 토대로 AFS 처리를 위한 Data Packet의 처리 절차와 히스토리 데이터베이스에 저장되는 전문의 형식을 나타낸 것이다.

Request ※1은 eGate로 카드 승인을 요청하기 위한 기본자료로서 Field의 내용은 Encrypt Module에 의해 암호화되어 전송된다. Response ※2는 카드 승인 결과를 상점으로 Return 되고 Field의 내용은 역시 Encrypt Module에 의해 암호화되어 전송된다. 그림에 따른 #1-6의 Data Packet의 처리 내용을 살펴보면 아래 표와 같

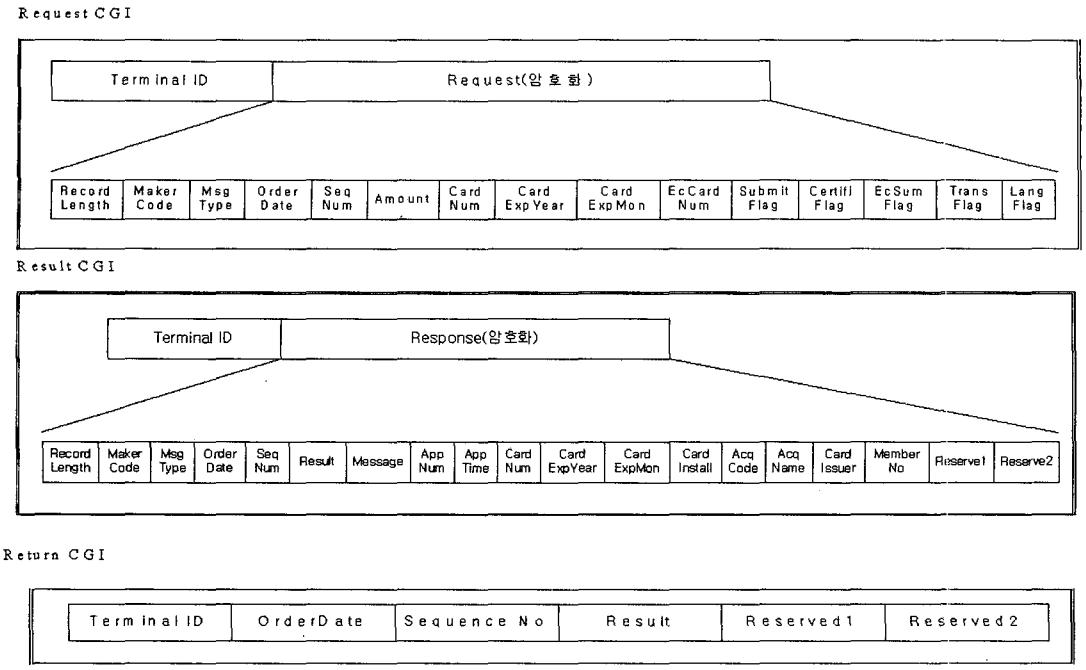

그림 7. AFS 처리를 위한 암호화된 전문의 형식
Fig. 7. Encryption Letters Form for AFS Processing.

표 1. AFS 처리를 위한 Data Packet의 처리 내용
Table 1. Data Packet Process Content for AFS.

ID	Data Packet	Etc
#1	TerminalID, Request(※1)	거래 정보
#2	카드정보	eGate 화면 기본제공
#3	TerminalID, Response(※2)	처리결과 정보
#4	Ack 전송	TRUE/FALSE
#5	처리결과 HTML	eGate화면(수정불가)
#6	Return CGI	종료 후 이동할 CGI

이 설명 할 수가 있다.

다음 그림 8은 주문되는 사기접속에 대해 제안된 Hybrid 인증으로 즉시 방어 처리하는 과정을 보여준다.

V. 결 론

기존 AVS의 카드 거래 승인을 위한 Payment Gateway의 기술적 한계점은 결제 처리에 따른 공유된 Database에 데이터마이닝 검증을 거치지 않는데 있다. 검증을 거치더라도 추출된 위험점수만을 신뢰하는 것이거나, 주소와 ZIP Code의 첫 번째 세 자리 숫자만을 비교하는 매개변수에만 크게 의존하는 분리된 방법이었다.

본 논문에서는 이러한 문제점을 효과적으로 해결하기 위한 방법으로 데이터마이닝의 적용과 순차적 AFS (Anti-Fraud System) 의 적용으로 기존의 AVS 시스템

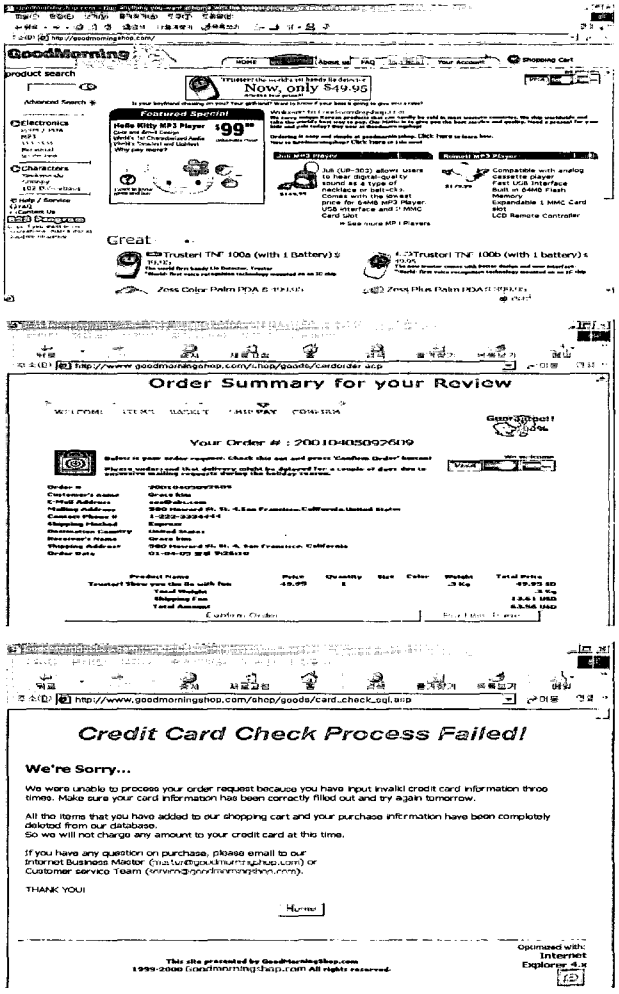

그림 8. Hybrid 인증 처리의 예
Fig. 8. Sample Authentication Form for Hybrid Proessing.

만으로 판단하기 어려운 사기 거래를 카드 인증과 동시에 사기 방어 막을 친다는 의미로서 고안된 시스템이다. 특히 Billing Address까지 해킹된 정보일 경우 본 시스템은 매우 효과적인 사기방지책이라고 할 수 있다.

본 논문에서 주장하는 AFS는 공유된 데이터베이스에 연관규칙 알고리즘 적용으로 1차로 사기율이 일정 %이상 일 때 걸러내고, 필요시 2차로 순차적 검정법을 적용하여 입력된 고객의 정보가 히스토리 정보와 일치하는지 여부를 분석 해낼 수 있으며 한계치 내에서 우량 고객정보가 입력되었을땐 연관규칙 알고리즘과 함께 빠른 인증허가가 이루어진다. 본 시스템에서는 히스토리 데이터베이스에 기록된 자료의 특성상 긴 연관규칙보다는 실용적인 면에서 길이가 2 내지 3인 연관규칙을 찾아 이러한 문제점을 극복하는데 초점을 맞추었다.

마지막으로 본 시스템은 히스토리 데이터베이스를 사용하기 때문에 은행 망에 연결할 필요성을 적게하며 새로운 결제가 이루어질 때마다 더욱 정확도가 개선되기 때문에 앞으로 시간을 두고 더 많은 사용자의 결제 패턴과 세계각지의 주소들을 효과적으로 정리한다면 전 세계적으로 통용될 수 있는 강력한 신용 카드 사기차단 시스템의 완성을 이룩할 수 있을 것이다.

참 고 문 헌

[1] Glasheen, C. and Dowling, S., "Increasing Internet Sales", Comm. IDC, Bulletin #W25213 - July 2001, Internet URL <http://www.idc.com/>.

[2] Information Technology OSI Systems Manageme -nt., Objects and Attributes for Access Control, ISO/IEC 10164-9, JTC1, 1995.

[3] P.K. Chan, and Fan, W., "Distributed Data Mi -ning in Credit Card Fraud Detection", IEEE Intelligent Systems, November/December 1999, pp. 67-74.

[4] SAS Institute Inc., Using Data Mining Techni -ques for Fraud Detection: A Best Practices Approach to Government Technology Solutions, Internet URL <http://www.sas.com/solutions/public_sector/white_papers/24596_0699.pdf/>.

[5] Provost, F. and Fawcett, T., "Robust Classifica -tion for Imprecise Environments. In: Machine Learning", Vol. 42, No. 3, pp. 203-231, 2001.

[6] Stolfo, S. Fan, W. Lee, W. Prodromidis, A. and Chan, P., "Cost-Based Modeling for Fraud and Instruction Detection: Results from the JAM Project", Proc. DARPA Information Survivability Conference and Exposition, IEEE Computer Press, pp.130-144, 2000.

[7] Merchant Works., E-commerce Glossary, Inter -net URL <http://www.merchantworkz.com/glossary.asp/>.

[8] Verisign White Paper., Alternative Approaches to Managing Fraud, Internet URL <http://www.verisign.com/rsc/wp/fraudscreen/fraudscreen_wp.pdf/>.

[9] 송용욱, 성기윤, "인터넷상의 전자 지불 시스템", Biz-on-Net, pp. 298-299, Internet URL <http://www.fv.com/>.

[10] Master Card International's., The Developments of the SET Protocol, Internet URL < http://www.mastercard.com/set/>.

[11] J.S. Park, M. Chen, and P.S. Yu., "Using a Hash-Based Method with Transaction Trimming for Mining Association Rules", IEEE Transac -tions on Knowledge and Data Engineering, Vol. 9, No. 5, pp.813-825, Sept. 1997. <http://www.sas.com/solutions/public_sector/white_papers/24596_0699.pdf/>.

[12] J. Han, and Kamber, M., Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, 2000.

[13] Agrawal, R. Mannila, H. Srikant, R. Toivonen, H. and A.I. Verkamo, "Fast Discovery of Asso -ciation Rules", In Advances in Knowledge Discovery and Data Mining, U.M. Fayyad, Piatetsky-Shapiro, G. Smith, P. and Uthurusamy, R. ed., AAAI Press/The MIT Press, pp. 307-328, 1996.

[14] Bonchi, F. Giannotti, F. Mainetto, G. and Pedreschi, D., "A classification-based Methodo -logy for Planning Audit Strategies in Fraud Detection", ACM Press New York, NY, USA, pp. 175 - 184, Series-Proceeding-Article Year of Publication, 1999.

[15] Tse-Hua, Lan. and Ahmed, H., "Fraud Detection and Self Embedding", ACM Press New York, NY, USA, pp. 33 - 36, Series-Proceeding-Article Year of Publication, 1999.

[16] Rosset, S. Murad, U. Neumann, E. Idan, Y. and Pinkas, G., "Discovery of Fraud Rules for Teleco -mmunications-Challenges and Solutions", ACM Press New York, USA, pp 409 - 413, Series -Proceeding-Article Year of Publication, 1999.

[17] Soheila, E., The Enhancement of Credit Card Fraud Detection Systems Using Machine Lear -ning Methodology, University Of Toronto (Canada), Mai, 38/06, p. 1640, Dec 2000.

[18] Lisa, M. Sas, K., No. 82 Aid Auditors in Finan

-cial Statement Fraud Detection?, University Of Colorado Boulder, DAI-A 58/07, p. 2732, Jan 1998.
[19] Charles, C. Steve, S., "A Critical Examination of Reengineered Audit Processes and The Likeliho -od of Detecting Fraud", No.3, pp. 297-310, 2002.
[20] Rosset, S. Murad, U. Neumann, E. Idan, Y. and Pinkas, G., "Discovery of Fraud Rules for Teleco -mmunications-Challenges and Solutions", ACM Press, New York, USA, pp. 409 - 413, Series -Proceeding-Article, Year of Publication, 1999.

저 자 소 개

조 문 배(정회원)
1988년 프랑스 INPL
제어계측학과 졸업(석사)
1998년 경북대학교 전자공학과
박사과정 수료
<주관심분야:신용카드보안통신, 전자상거래기술>

박 길 흠(정회원)
제40권 제4호 참조