

A 은행 사례 분석을 통한 빅데이터 기반 자금세탁방지 시스템 설계

A Study on Big Data Anti-Money Laundering Systems Design through A Bank's Case Analysis

김상완¹ · 함유근^{2†}

(주) 바넷정보기술¹
건국대학교 경영대학²

요약

본 논문은 타 산업에 비해 상대적으로 빅데이터의 잠재 가치가 높은 금융권에서 사전 예방과 실시간 탐지가 특히 중요해지고 있는 자금세탁방지를 위한 빅데이터 기반의 시스템 구현 방안을 제시한다. 기존 A 은행의 사례 분석을 통해 현 자금세탁방지 업무에 빅데이터를 적용 가능한 부분과 기능들 논의한다. 연구 결과 빅데이터를 활용한 자금세탁방지 시스템은 기존의 자금세탁방지 시스템에 추가적으로 소셜 네트워크 서비스나 로그 파일 등의 다양한 출처로부터 데이터를 수집해 일부가 아닌 전체 데이터를 대상으로 빠른 속도로 거래 모니터링과 잠재적 위험 탐지가 가능하다. 즉 빅데이터 기반 자금세탁방지 시스템은 비정형 데이터인 소셜 네트워크 서비스 데이터와 기존 거래나 고객정보를 통합하고 소셜 네트워크 분석 방법을 적용하여 고객확인업무와 잠재 요주의 인물 탐지 기능을 제고시킬 수 있다. 또한 빅데이터 분석 인프라 하에서 룰 모델이나 거래패턴 스코어링 모델을 적용해 실시간으로 혐의 거래 적발이 가능하다.

■ 중심어 : 빅데이터, 자금세탁방지 시스템

Abstract

Traditional Anti-Money Laundering (AML) software applications monitor bank customer transactions on a daily basis using customer historical information and account profile data to provide a “whole picture” to bank management. With the advent of Big Data, these applications could be benefited from size, variety, and speed of unstructured data, which have not been used in AML applications before. This study analyses the weaknesses of a bank's current AML systems and proposes an AML systems taking advantage of Big Data. For example, early warning of AML risk can be improved by exposing identities and uncovering hidden relationships through predictive and entity analytics on real-time and outside data such as SNS data.

■ Keyword : Big Data, Anti-Money Laundering Systems

I. 서론

2011년 우리나라 관세청의 불법 외환거래 적발 실적은 약 3조 8천억 원으로 2007년 약 2조 3천억 원에 비해 4년 만에 약 61.3%가 증가했고 향후 연간 5조 원 이상을 기록할 것으로 예상되고 있다. 하지만 실제 자금세탁을 포함한 불법 외환거래 수법은 갈수록 복잡해

지고 있어 발생 건수 대비 적발 건수의 비율은 5%에 불과하다. 최근 범죄조직의 경제력 증대로 인해 첨단기술 및 금융전문가들을 활용한 고도의 자금세탁기술이 지속적으로 개발되고 있다. 또한 범죄조직에 의한 금융회사의 경영권 인수 및 정치권과의 결탁도 증대되고 있다. 그리고 다양한 금융기법을 통해 합법자금과 불법자금을 섞어서 전 세계로 자금을 분산하여 사법당국의

범망을 교묘히 피해가고 있다[4]. 이러한 범죄를 해결하기 위해 자금세탁방지법이 있으며 은행들은 자금세탁방지 시스템을 이용해 관련 위험을 분석, 탐지하고 있다. 결국 이렇게 날로 전문화되고 있는 자금세탁 거래를 방지하기 위해서는 더욱 고도화된 해결책이 요구되고 있다. 그런 새로운 해결책으로 최근 빅데이터의 활용이 주목을 받고 있다.

빅데이터는 금융 서비스 회사들에게 특히 유망하고 차별화 요소를 가지고 있다. 금융 서비스 회사들의 상품은 물리적인 제품의 형태를 가지고 있지 않다. 서비스에 활용하고 있는 정보의 원천인 데이터는 가장 중요한 자산이다. 금융회사들은 비즈니스와 금융관리에 이용되는 거래 데이터(Transaction Data)는 그동안 여러 해 동안 축적해 왔다. 하지만 빅데이터 시대로 진입 하면서 매일 축적되는 데이터의 양이 과거보다 더욱 폭발적으로 증가하고 있다. 이로 인해 금융회사들은 다른 산업보다 더 많이 보유하고 있는 데이터를 활용하여 경쟁우위를 얻을 가능성이

높아지고 있다. 2012년 미국 옥스퍼드 대학교(University of Oxford)와 IBM이 빅데이터의 실제 사용에 관한 공동 조사한 연구 보고서에 의하면 은행 및 금융회사의 71%는 빅데이터의 활용이 자신의 조직에 대한 경쟁력을 높인다고 주장한다[11].

빅데이터를 활용하여 비즈니스에 반영하는 해외 금융회사의 사례는 많지만 아직까지 국내 금융회사들이 자금세탁방지를 위해 빅데이터를 적용한 사례가 미비하다. 본 연구에서는 타 산업에 비해 상대적으로 데이터 보유량이 많고 빅데이터 잠재가치가 높은 금융권에서 적용 가능한 빅데이터 기반의 자금세탁방지 비즈니스 인텔리전스에 주목하였다. 현재 일부 카드사와 보험사가 이상패턴 검출과 부정사용 예방의 목적으로 빅데이터를 적용하고 있다. 이러한 부정사용 감시의 목적이 자금세탁방지(Anti-Money Laundering)¹⁾와 유사한 부분이 많다.

자금세탁방지 시스템은 최근 국내/외 범죄조직의 경제력 증대로 인해 첨단기술과 금융 전문가를 활용한 고도의 자금세탁 기술에 대응해야 하는 과제를 가지고 있다. 본 연구는 기존 자금세탁방지 시스템의 문제점을 분석하고, 빅데이터를 활용해 이를 극복할 수 있는 시스

템을 제시한다. 특히 자금세탁방지 시스템에서 빅데이터를 어떻게 활용할 수 있을지 시스템적인 측면에서 구현하는 것을 목적으로 한다.

II. 관련 연구

2.1 금융 빅데이터

2001년 미국의 시장조사 및 연구 컨설팅 회사 가트너(Gartner)의 더그 레이니(Doug Laney)는 빅데이터의 특징을 규모(volume), 다양성(variety), 속도(velocity) 등 '3V'로 불리는 3가지 핵심요소로 규정했다. 빅데이터는 업무 수행 방식 측면에서는 다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고, 데이터의 빠른 수집(High-velocity Capture), 발굴(Discovery), 분석을 지원하도록 고안된 차세대 기술 및 아키텍처라고도 한다[9].

특히 금융산업에서는 이러한 3V의 특징을 가진 빅데이터가 의사결정 과정에서 매우 가치가 높은 것으로 평가받고 있다[12]. 금융회사들에게 빅데이터를 활용하기에 유용한 원천 데이터 포맷은 거래에 관련된 트랜잭션 및 로그 데이터이다. 이에 추가해 은행 비즈니스 또는 정보시스템 운영의 세부 사항을 기록하기 위해 생성해 왔지만 분석되지 않은 거래 데이터, 로그 데이터, 이벤트 데이터와 이메일 내용 등의 비정형 데이터들도 활용 가능성이 타 산업에 비해 높다[11]. 복잡하고 다양한 유형의 구조화/비구조화 데이터를 취급하는 은행에 있어서 운영 효율성의 개선은 성공을 위한 필요 요소라고 할 수 있다. 데이터를 수집하는 즉시 접근하고 분석하는 능력은 비즈니스 전반에 걸쳐 적시에 고도의 분석적 통찰을 얻어내기 위한 기초가 된다. 또한 빅데이터 분석을 통해 기존 IT 리소스(인력, 자금, 인프라)의 활용을 극대화하는데 필요한 확장성을 얻을 수 있고 효율성을 보장할 수 있다[7].

2.2 자금세탁방지 및 기술

자금세탁(Money Laundering)²⁾이란 일반적으로 마약

1) 국내외로 이루어지는 불법 자금 세탁을 적발하고 예방하기 위한 법적·제도적 장치.

2) 1920년대 미국에서 '알카포네'와 같은 조직범죄자들이 도박이나 불법 주류판매를 통한 수입금을 자신들의 영향력 아래에 있고, 현금거래가 빈번한 세탁소(Laundry)의 합법적 수입으로 가장한 것으로부터 유래되었다.

거래, 조직폭력 등과 같은 범죄 행위로부터 얻은 불법적 수익에 대한 자금출처를 은닉함으로써 합법적인 자산으로 위장하는 과정을 의미한다. 무기거래, 밀수, 마약 거래, 매춘 등과 같이 조직 범죄활동을 통해 형성된 불법자금이나 사기, 뇌물, 횡령, 정치적 커미션 등에 의해서 만들어진 불법자금을 합법화하기 위한 수단으로 자금세탁이 활용되고 있다[4].

국제 자금세탁방지 기구인 FATF(Financial Action Task Force on Money Laundering)³⁾는 자금세탁행위를 ‘범죄수익의 불법원천을 가장하기 위한 과정’으로 정의하고 있다. 우리나라는 1993년 금융실명제를 도입하면서 자금세탁이란 용어를 활발하게 사용하게 되었고, 외국과는 다르게 탈세목적으로 외국환 거래를 이용하는 경우를 포함하고 있는 것이 특징이다[3].

자금세탁 행위자들의 목적은 자금의 원천을 위장하고 더러운 돈(Dirty Money)을 세탁(Wash it)하여 소득원천의 재 추적이 불가능한 형태로 만드는 것이다. 자금세탁의 대상이 되는 자금을 은행 계좌로 예치하거나 부동산, 주식, 보험납입금 및 기타 자산으로 배치(Plac-

ing)하여 향후에 의심을 받지 않고 사용할 수 있도록 한다. 불법자금을 명백한 합법자금 또는 자산으로 전환하는 자금세탁의 기본과정은 국제적으로 미국 관세청 3단계 모델 이론에 따라 배치(Placement)-반복(Layering)-통합(Integration)으로 구분할 수 있다[6].

자금세탁 방지를 위한 거래 모니터링 업무는 고객 및 거래 데이터를 주기적으로 분석하여 자금세탁 행위로 의심되는 금융거래의 추적을 원활히 지원하는 역할을 한다. <표 1>과 같이 효과적인 이상거래 추적을 위하여 감시리스트(Watch List) 필터링, 룰 모델, 스코어링 모델, 이상거래 패턴분석 모델 등의 다양한 분석기법이 거래모니터링 시스템에 활용되고 있다[4]. 특히 인터넷 등 디지털화와 국제 거래의 증가에 따른 거래 데이터의 폭증으로 데이터마이닝 등 컴퓨터에 의한 자동 탐지 기술의 활용도가 높아지고 있다[10]. 이런 기술들은 이미 해외에서는 2000년대 초반부터 사용되고 있다[8].

III. 연구방법 및 사례 분석

3.1 연구 방법

본 연구는 사례 분석 방법을 이용해 기존 시스템의 문제점을 찾아낸 후 이를 극복하기 위한 빅데이터 적용 방안을 탐구하였다. 이를 위해 먼저 기존 국내 A은행 자금세탁방지 시스템의 문제점과 개선방향을 심층 사례분석을 통해 진행하였다. 사례 분석의 정합성을 높이기 위해서는 다양한 여러 관련 사례들을 함께 분석해야 하지만 적절한 국내 사례의 부재로 한 사례에 대한 심층 분석에 중점을 두었다. 그리고 사례 분석 과정에서 보험산업의 사기 탐지와 같은 타산업 사례들을 참고하였다.

3.2 A 은행 자금세탁방지 시스템

A 은행은 급변하는 금융환경의 변화에 따라 고객 및 거래의 다양한 위험요소가 증가하고 있다. 더욱이, 규제 중심의 거래 및 고객 모니터링에서 벗어나 환경 변화에 능동적으로 대처하고 효율적 관리 체계로 전환할 필요성도 커지면서 자금세탁방지 시스템을 구축하였다.

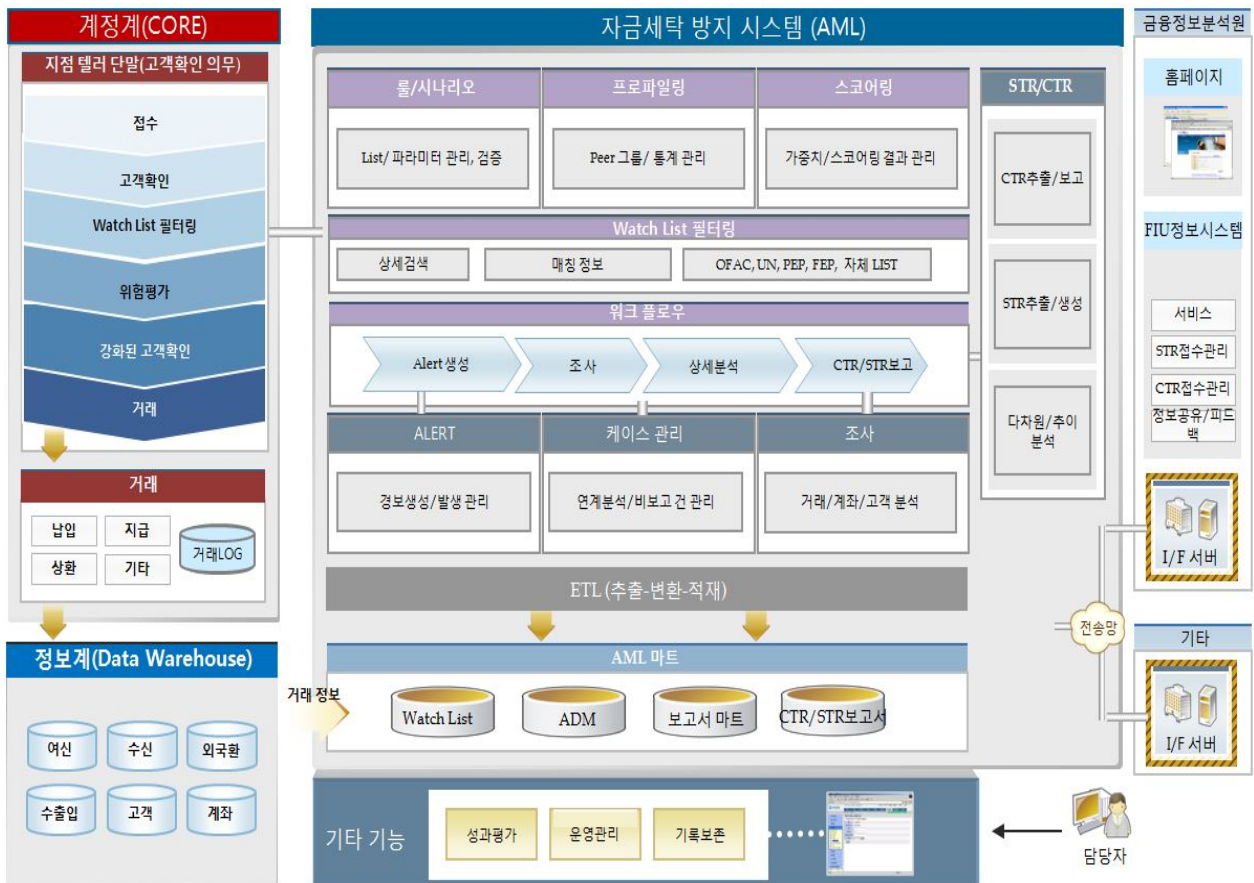
자금세탁방지 시스템은 <그림 1>과 같이 영업점에서 고객정보를 획득하거나 획득된 정보를 주주의 고객리스트와 매칭하고 필터링(Watch List Filtering)하여 확

<표 1> 자금세탁방지 거래 모니터링 분석기법

주요 기술	내용
룰 모델 (Rule-Based Model)	<ul style="list-style-type: none"> 일정금액 이상, 일정기간 분할거래 횟수 등 특정 거래 또는 행위 기반으로 의심거래 추출 금융거래 조건 및 환경 등이 변함에 따라 지속적인 업데이트 필요
스코어링 모델 (Scoring Model)	<ul style="list-style-type: none"> 거래, 고객, 국가 등 자금세탁 위험 요인을 고려하여 스코어링을 수행하고 스코어링 결과에 따라 위험을 파악
이상거래 패턴분석 (Outlier Detection)	<ul style="list-style-type: none"> 나이, 직업 등 고객의 성격에 따라 유사그룹을 분류한 후 고객 및 그룹별 프로파일을 생성하고 이후 거래 발생 시 분석대상 고객과 유사그룹 프로파일을 비교하여 이상거래 탐지 거래패턴 인식을 통해 알려지거나 또는 알려지지 않은 이상거래 탐지 가능
Watch List 필터링 (Watch List Filtering)	<ul style="list-style-type: none"> 금융회사가 고객을 대상으로 고객의 이름과 테러리스트 명단, 범죄자 명단을 비교 분석

출처: 삼일회계법인, 2013.

3) 대표적인 선진국 자금세탁방지 기구로서, 1989년 파리에서 개최된 G7정상회의에서 자금세탁방지외 이에 대한 국제공조의 중요성을 인식한 정상들의 지시에 따라 설립되었다.



출처: SAS, 삼성SDS 자료 재구성.

〈그림 1〉 A 은행 자금세탁방지 시스템 개념 구성도

인업무를 수행하는 위험기반 고객확인 부분이 있다. 그리고 고객이나 상품 등의 위험관련 분석정보를 관리하는 프로파일링 정보와 룰(Rule)이나 스코어링(Scoring)을 이용하여 거래 위험의 측정과 자금세탁 혐의거래를 검출하는 거래 모니터링(Transaction Monitoring) 부분이 있다. 또한 혐의거래 및 혐의고객 계좌 개설에 대한 분석과 사례를 관리하는 사례분석 관리(Case Management) 부분도 있다. 마지막으로 일정기준 이상의 혐의 등급을 가진 거래를 감독기관인 금융정보분석원에 보고하는 대내/외 리포팅 부분이 있다[2].

자금세탁방지 시스템의 대략적인 기능은 다음과 같다. 위험기반 고객확인 은 일반적으로 금융거래를 원하는 고객의 자금세탁위험을 평가하기 위해 크게 고객(Customers), 서비스(Services), 상품(Products), 국가(Geographies)로 구분하여 각 위험 별로 점수를 산정하여 고객위험을 평가한다. 그리고 위험등급에 따라 검토하거나 요주의 고객 필터링(Watch List Filtering)을 통해 국내나 국외에 자금세탁의 위험이 높은 요주의 인물 목록

정보를 만든다. 목록을 활용하여 신규 계좌 개설의 경우나 의심스러운 고객을 분석 대상으로 선정하거나 거래를 거절할 수 있는 근거로 활용한다.

거래 모니터링은 의심스러운 거래를 적발하는 기능이다. 적벌 과정에 혐의거래 추출의 기초분석에 사용되는 거래 패턴 분석 모델이나 연계 분석 모델과 같은 룰베이스 모델 및 위험 스코어링 모델 등이 사용된다. 룰베이스 모델은 특정한 추출기준을 설정하고 해당 기준에 부합하는 거래를 추출하여 분석하는 모델이다, 반면 스코어링 모델은 거래특성 변수 별로 가중치(Weight)를 부여하여 해당 거래의 혐의 정도를 계량화하여 수치로 표현하는 모델이다. 거래 모니터링에서 일정한 혐의 등급 이상을 부여 받은 거래정보는 경고(Alert)로 표시되어 심사분석 단계로 배당된다.

사례분석 관리는 영업점이나 거래모니터링 시스템으로부터 자금세탁의 혐의가 의심되는 고객이나 거래가 보고되면 담당자 지정이나 배당을 거쳐 상세분석 등 일련의 과정을 처리하는 심층 분석업무 지원기능이

수행된다. 협의도가 낮은 거래는 필요 시 추후 활용될 수 있도록 데이터베이스에 저장하여 보존 처리한다. 이들 정보가 추후에 발생하는 신규 보고와 연관성이 있을 경우 상세분석을 할 수 있다. 협의도가 높은 거래의 경우는 자동배당기능(Assignment & Routing)을 활용하여 분석자가 상세하게 분석할 수 있다. 이와 같이 위험도에 따른 차등 처리로 보다 많은 거래의 분석이 가능하고 분석 품질이 높아지는 등 자금세탁방지의 효율성을 제고할 수 있다.

마지막으로 대내/외 리포팅 기능은 감독기관이 요구하는 형식에 맞게 보고서를 자동으로 작성하고 전송하며 불필요한 문서 작성 및 보고 시간을 줄여준다. 그리고 내부적으로는 자금세탁방지 업무의 성과 관리를 위한 통계자료나 기존 혐의거래 보고서와의 내용 조화까지 한 화면에서 통합적으로 관리할 수 있다.

3.3 정보 처리 측면의 개선점

3.3.1 정보 수집

정보 분석의 질을 높이기 위해서는 정보 수집이 무엇보다 중요한 활동이다. 금융회사가 자금세탁방지를 위해 수집해야 하는 정보는 금융회사 내부의 거래정보, 금융위원회 또는 OFAC(미국 재무부 해외자산통제국) 등에서 제공하는 금융거래 제한 대상자 리스트에 오른 내/외국의 정치적 주요인물, 공중협박자금⁴⁾ 조달 고객, 종합자산관리 서비스⁵⁾ 대상인 고객의 정보 등이다. 그런데 이들 중 특히 금융거래 제한 대상자 리스트는 현재 자금세탁방지 시스템에서 배치로 작업하고 있다. 더욱이 관계 기관으로부터 받은 자료를 자동으로 갱신하지만 자료를 제공하는 솔루션 업체가 월드체크(World-Check)사나 팩티브(Factivs)등의 외국계 회사이기 때문에 기업의 비자금이나 정치자금 조성을 통해 자금세탁 행위가 주를 이루어지는 국내 실정을 완벽하게 반영하기는 어렵다.

금융회사 내부에 보유하고 있는 거래내역 데이터나 자체 고객정보 데이터는 추출-변환-적재 작업을 통해 데이터 웨어하우스 시스템으로 데이터를 이관하여 자

금세탁방지 데이터 마트(Data Mart)에 반영하고 있다. 이러한 데이터 마트는 금융정보분석원으로 보고하기 위한 목적으로 데이터가 재가공 되었기 때문에 실제 자금세탁의 기법을 파악하거나 분석시스템으로 활용하기에는 한계가 있다. 따라서 내/외부의 다양한 정보를 수집하여 분석에 활용할 수 있는 별도의 데이터베이스나 구조가 필요하다.

거래 모니터링의 경우에도 최근 채널의 다양화로 금융거래가 전자이체 등의 비대면 거래가 많아지고 비계좌 거래나 신용카드, 현금카드 또는 상품권이나 채권, 유가증권 등의 지급수단이 다양화 되면서 갈수록 어려워지고 있다. 즉, 거래 자체가 지능화되고 복잡해지는 혐의의심 거래에 대한 분석을 내부의 거래정보의 데이터만을 활용하여 판단하기에는 역부족이다. 따라서 외부 기관과 연계한 다양한 데이터의 원천으로부터 실시간으로 빠르고 정확하게 이벤트나 정보를 수집하여 시스템에 활용할 수 있게 실시간 데이터 수집 기능을 강화해야 한다.

3.3.2 정보 분석

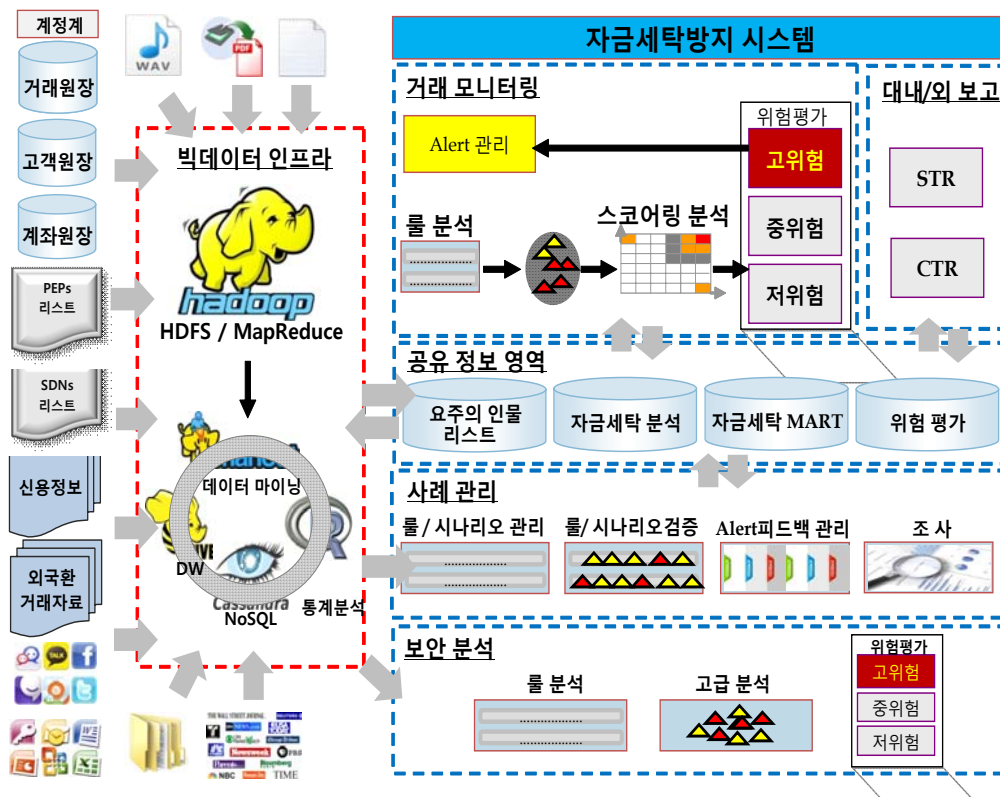
금융회사들이 금융정보분석원에 보고한 의심거래보고 건수는 해마다 증가하는데 비해 보고내용의 질적인 측면에서는 크게 개선되지 않았다. 지난 최근 3년간 금융정보분석원의 심사분석/제공 현황을 보면 의심거래보고 대비 실제 제공비율(법집행 기관 제공건수/의심거래 보고 건수)은 약 6% 수준으로 오히려 보고 품질은 하락하였다[1, 4].

의심거래보고 건수 및 고액현금거래보고 건수의 급증과 고도화 및 국제화되고 있는 자금세탁 행위에 적극 대응하기 위해서는 고액현금거래 보고가 의심거래 보고 분석의 보조 자료로만 사용되어 오던 부분의 개선이 필요하다. 그리고 의심거래 보고와 고액현금거래 보고의 연계 분석을 강화하여 차명거래 개연성이 높은 계좌를 찾아내거나 악성 분산송금, 기업의 비자금 등 자금세탁행위 적발 역량을 높일 필요가 있다.

앞서 언급했듯이 자금세탁방지 거래 모니터링 분석에는 룰 베이스 분석 모델이나 스코어링 분석 모델이 사용된다. 그런데 자금세탁으로 의심되는 거래의 선별이 실시간으로 분석되지 않아 모델 분석과 검증에 일정시간이 소요된다. 또한 전체가 아닌 일부의 샘플 데이터를 통해 모델을 설계해 업무에 이용하기 때문에 신규 혐의거래 유형을 즉각적으로 시스템에 반영하기 어렵다.

4) 국가나 지방자치단체 또는 외국정부의 권한행사를 방해할 목적으로 공중에게 위해를 가하고자 하는 등의 공중을 협박할 목적으로 사용하기 위해 사용되는 자금이나 재산.

5) 매분기 직전 12개월의 기간을 기준으로 1년 평균 10억 원 이상(비거주자 25만 불)의 부유층 고객을 대상으로 투자자문을 비롯한 법률, 세무 설계 등 종합적인 자산을 관리하는 서비스.



〈그림 2〉 제안된 빅데이터 자금세탁방지 시스템 개념 구성도

다양한 정보를 활용하여 의심스러운 거래보고를 분석하는데 링크 분석(Link Analysis)⁶⁾과 같은 선진 분석 기법의 사용하여 고객정보나 거래정보를 기반으로 여러 계좌간의 숨겨진 연관 관계를 확인할 필요가 있다. 예를 들어 텍스트 마이닝 기술을 활용하여 비정형화 되어 있는 거래로그나 혐의거래 분석에 사용된 다양한 텍스트 문서에서 보고 사유에 대한 유형을 분류할 수도 있다. 또한 자금세탁 혐의고객의 주변 관련자에 대한 정보도 추출이 필요하다. 이런 과정에서 다양한 데이터 소스와 연동하여 비주얼한 패턴 분석과 즉각적인 연관분석을 쉽게 도와주는 시각화 솔루션의 사용도 필요하다.

3.3.3 정보 관리 및 공유

영업점 고객확인이나 감사 주관부서의 의심스러운 거래 심사분석에 사용되는 자료는 계좌개설 관련 서류 또는 실명확인표 사본, 검증자료 원본 등이다. 해당 자료는 서면으로 주고받거나 비문서적 방법인 전자정부 홈페이지를 통한 주민등록증 진위확인, 신용정보기관 확

인방법, 실제 본인 음성 녹취 등 다양한 방식으로 이루어지고 있다. 때문에 심사분석 자료를 안전하고 적절한 시기에 정보가 사용되도록 제공 받는 데는 어려움이 있다.

고객확인을 실시하는 과정에서 발생하는 검증자료나 금융거래기록, 의심스러운 거래보고서를 포함한 내/외부 보고서 및 관련 자료는 5년 이상 보안이 유지되도록 자료를 관리해야 하는 어려움이 있다. 금융정보분석원과의 정보 공유는 서면으로 주고받는 과정에서 발생하는 정보의 유출 위험을 방지하기 위해 전자상거래에서 적용하는 암호화, 전자서명 등의 최신 기능을 이용하여 온라인으로 현재 이루어지고 있으나 금융회사 간에는 정보 공유가 직접 이루어지고 있지 않다.

전산화 형식으로 보관하지 않은 고객확인 의무 관련 자료는 영업점에서 보존관리하고 의심스러운 거래보고 자료는 주관부서나 보고와 관련된 영업점에서 보존하고 있다. 따라서 보존된 자료들을 금융정보분석원에 제공하거나 해당 자료를 추후 고객확인 의무 이행이나 자금세탁 혐의거래 분석에 즉시 활용하기 어렵다. 고객확인 의무나 혐의 심사분석에 사용된 자료들을 통합해서 관리해 쉽게 공유될 수 있다면 추후에 추가적인 고객확인 의무 수행 시나 거래모니터링의 분석 모델에서

6) 시간의 흐름에 따른 구매 유형이나 개별 부문들 간 구매유형간의 상관관계를 알아보기 위한 데이터 마이닝 기법.

활용이 가능하다.

3.4 데이터 측면의 개선점

3.4.1 데이터 규모

기존 자금세탁방지 거래 모니터링에 사용하는 룰 모델이나 스코어링 모델은 제한된 데이터로부터 전체의 특성을 알아내는 표본 추출(Sampling)이나 통계 기법을 사용하여 모델을 설계하고 가중치를 산정하여 모델에 반영하고 있다. 하지만 이러한 분석 방법은 분석에 필요한 정형화된 형태로 데이터를 가공하여 사용해야 하기 때문에 계획적인 공모에 의한 신규 범죄행위의 거래 패턴을 찾아내기는 어렵다.

많은 데이터를 한꺼번에 처리할 수 있는 빅데이터의 규모 측면 특징을 자금세탁방지 시스템에 적용할 수 있다면 문제를 개선할 수 있다. 표본 추출이 아닌 데이터의 전수 분석을 통해 혐의거래 위험평가 모델의 개발이나 시나리오 테스트 과정에서 발생하는 정보의 왜곡을 줄일 수 있고 그동안 소실되었던 정보의 가치에 주목할 수 있다. 또한 막대한 데이터를 바탕으로 링크 분석과 같은 새로운 데이터 분석 기법을 적용하여 다양한 분석이 가능하다[5].

3.4.2 데이터 다양성

최근 채널의 다양화로 비대면 금융거래가 많아지고 지급수단이 늘어나게 되면서 이에 관련된 이벤트 로그 정보들이 쏟아지고 있다. 이러한 비정형 형태의 이벤트 로그 정보와 외부 기관의 다양한 자료까지 거래 모니터링이나 고객확인 의무 과정에 사용할 수 있다면 혐의거래 적발의 정확도를 높일 수 있다.

그동안 거래 모니터링 분석에 사용되는 데이터는 내부 거래내역 데이터, 자체 고객정보와 외부 기관에서 제공하고 있는 금융거래 제한 대상자 관련 정보가 전부다. 하지만 빅데이터 기반으로 자금세탁방지 시스템을 활용한다면 기존에 거래 데이터와 고객 데이터를 소셜 네트워크 서비스 데이터나 신용정보 회사와 같은 다양한 기관으로부터 수집된 비정형 데이터와 결합이 가능하다. 정형 데이터와 비정형 데이터를 바탕으로 분석하여 관리하는 정보들은 자금세탁방지 고객확인 의무 과정에서 확인하는 요주의 인물 리스트 필터링이나 자금세탁 위험평가 모델의 위험요인 분석에 활용도가 높다.

3.4.3 데이터 속도

대량의 데이터를 실시간으로 빠르게 처리할 수 있는 빅데이터의 특징은 자금세탁거래 모니터링에 사용하는 위험평가 모델의 개발시간이나 시나리오 테스트 시간을 단축시킬 수 있다. 지금까지 위험평가 모델의 개발은 비용과 시간의 문제로 인해 전체가 아닌 일부 데이터를 모집단으로 선정하여 가중치를 조절하면서 오랜 시간 동안 여러 번 반복적으로 시나리오 테스트를 수행했다. 빅데이터 환경에서는 모델 개발 시 전체 데이터를 대상으로 모델을 설계하고 빠르게 검증할 수 있기 때문에 실시간에 가까운 거래 모니터링이 가능하다. 실시간으로 빠르게 거래 모니터링이 가능해지면 자금세탁 혐의 분석과 적발 시간이 단축되기 때문에 혐의 거래에 대한 신속한 보고도 가능해 진다.

IV. 빅데이터 자금세탁방지 시스템 구현

빅데이터 자금세탁방지 시스템의 구성 요소로는 첫째, SNS 데이터와 같은 모바일이나 인터넷 상의 비정형 데이터를 통해 잠재적 위험과 관련된 잠재적 요주의 인물을 찾는 기능이 가능하다. 둘째, 기존 외부로부터 제공 받는 요주의 인물 리스트들을 이메일, 뉴스 피드 등 하둡에 저장된 비정형 데이터를 반영한 매칭 스코어로 요주의 인물 리스트를 검증하는 기능이 가능하다. 셋째, 기존의 배치 방식 대신에 실시간으로 거래를 모니터링 하는 기능도 가능해지고 있다. 넷째, 다양한 자금세탁 사례들을 분석해 특징을 찾아내는 부분도 중요한 요소이다. 이들을 각각 좀 더 자세히 살펴보면 다음과 같다.

4.1 소셜 네트워크 분석을 활용한 잠재 요주의 인물 탐지

기존의 자금세탁방지 시스템을 포함한 금융회사의 사기방지 시스템, 위험 관리 시스템(Risk Management System) 등의 시스템들에서는 특정 개인에 대한 개별적인 접근법(Silo Approach)이 일반적이었다. 하지만 최근 소셜 네트워크 분석(Social Network Analysis, SNA)이 도입되면서 통합적인 접근(Integrated Approach), 즉 관련 인물들을 함께 분석하는 것이 가능해지고 있다. 네트워크상의 인물들 간의 관계나 패턴을 분석해 잠재적인 금융 범죄의 위험을 예방할 수 있다. 보다 구체적으로는 자금세탁방지 시스템에서도 소셜 네트워크 서

비스를 활용하여 자금세탁 혐의자들 상호간의 공모 가능성을 확률적으로 해석하여 혐의그룹을 추출하는 기법을 사용할 수 있다. 영업점에서 고객확인 의무를 수행할 경우 이러한 잠재 요주의 인물을 탐지할 수 있다면 자금세탁을 시도하는 거래를 사전 예방하는 효과를 볼 수 있다.

4.2 하둡 인프라 기반의 요주의 인물 리스트 관리

금융회사는 거래가 발생할 때 고객확인 의무를 수행하며 요주의 인물 리스트 정보를 비교하여 자금세탁의 위험을 확인을 해야 한다. 현재는 보통 외국계 기업으로부터 자금세탁 중요 인물에 대한 리스트를 받거나 FATF(국제자금세탁방지기구)나 UN과 같은 국제기구로부터 자금세탁 혐의자들에 대한 정보를 받아 요주의 인물 리스트 데이터베이스에서 관리하고 있다.

최근 금융거래는 채널의 다양화와 비대면 거래의 증가로 소셜 네트워크 서비스를 이용한 금융거래와 같은 신규 위험요소들이 발생하기 때문에 기존의 특정회사에서 제공하는 요주의 인물 리스트만으로는 복잡하고 다양해지는 혐의거래를 찾아내기에는 한계가 있다. 하지만 대용량 비정형 데이터의 분산 처리를 지원하는 하둡 인프라와 맵리듀스를 사용하여 그동안 데이터를 수집하는 방법 외에 추가적으로 신용정보, 외국환 거래 정보와 같은 반정형 데이터나 소셜 네트워크 서비스, 뉴스 미디어, 외부기관의 텍스트 문서파일, 이메일 등의 비정형 데이터를 기존 자금세탁 방지 마트의 정형 데이터와 통합하여 요주의 인물 리스트로 관리한다면 갈수록 지능화 되어가는 자금세탁 위험에 대해 즉각적인 대처가 가능하다.

4.3 실시간 빅데이터 거래 모니터링

지금까지는 자금세탁으로 의심되는 거래를 예측 분석하기 위해서는 사전에 분석용 데이터베이스에 분석이 쉽게 정형화된 테이블을 설계하고 일정기간의 거래 데이터를 활용하여 분석 모델을 만들고 검증하는 과정이 선행되어야 했다. 하지만 이러한 모델 개발 과정은 빨라도 며칠이 걸리는 작업으로 실시간으로 모니터링하여 예측 분석해야 하는 환경에는 적합하지 않다.

빅데이터 분석 인프라를 이용하여 분석할 수 있다면 이러한 문제점을 개선할 수 있다. 실시간으로 운영 데

이터베이스의 거래정보나 고객정보, 요주의 인물 리스트 정보, 다양한 형태의 대내/외 심사분석 자료를 하둡 인프라 기술을 이용하여 분산 파일 시스템에 저장하고 하이브(Hive)와 같은 빅데이터 데이터 웨어하우스나 NoSQL 데이터베이스를 통해 통계분석, 데이터 마이닝을 수행하여 분석 데이터베이스에 필요한 정보를 저장할 수 있다. 그동안 룰 분석 모델이나 스코어링 모델과 같은 거래 분석 모델은 일부 샘플 데이터를 통해 모델의 시나리오를 검증했지만, 빅데이터 분석 인프라 환경에서는 일부가 아닌 분산 저장된 모든 데이터를 활용하여 빠르게 시뮬레이션 할 수 있기 때문에 거래 모델의 신속성과 정확도를 높일 수 있다.

4.4 사례분석 관리

자금세탁 혐의거래에 관련하여 조사활동이나 보고 과정에서 생성된 검증자료나 금융거래기록, 의심스러운 거래 내용을 포함한 내/외부 보고서 및 관련 자료는 5년 이상 보안이 유지되도록 관리해야 한다. 그리고 필요에 따라서는 다시 조회하거나 보고할 수 있어야 한다. 하지만 기존의 시스템은 고객확인 의무를 수행하는 과정에서 발생하는 전산화 되지 않은 형식의 자료들은 관리하지 않기 때문에 추후에 자금세탁 혐의거래 분석에 자료들을 즉시 활용하기는 어렵다.

빅데이터를 활용하여 사례분석 관리를 한다면 혐의 관련 거래 데이터와 조사 분석 과정에서 사용된 검증자료나 음성 녹취 오디오, 주민등록증 진위확인 등 물론 신용정보기관 확인 과정에서 발생하는 데이터들을 추후에도 활용할 수 있다. 빅데이터 인프라로 처리하여 기존의 룰이나 시나리오와 맞는지 검증/관리 할 수 있다. 또한 사례분석과 관련하여 경고 피드백 내용은 추후에 조사 분석 담당자가 자금세탁방지 마트를 통하여 언제든지 확인이 가능하다.

V. 결론

본 연구에서는 데이터 보유량이 상대적으로 많고 빅데이터 잠재가치가 높은 금융권에서 활용이 가능한 분야로서 자금세탁방지 시스템 구축 방안을 제시하였다. 기존의 자금세탁방지 시스템은 자체적으로 관리하고 있는 고객 정보나 거래 정보의 샘플 데이터를 활용하

〈표 2〉 빅데이터 자금세탁방지 시스템의 특징

	기존 자금세탁방지 시스템	빅데이터 자금세탁방지 시스템
소스 데이터	<ul style="list-style-type: none"> ◦ 거래내역(여신, 수신, 외환) ◦ 고객정보(내부) ◦ 혐의 분석 정보(내부) ◦ PEPs 리스트(솔루션) ◦ OFAC 금융제한 대상자 정보 ◦ 금융위원회 거래제한 대상자 ◦ 자체 보고 자료(STR/CTR) 	<ul style="list-style-type: none"> ◦ 거래내역(여신, 수신, 외환) ◦ 내/외부 고객정보 ◦ 내/외부 혐의 분석 정보 ◦ PEPs 리스트(솔루션 or 자체) ◦ OFAC 금융제한 대상자 정보 ◦ 금융위원회 거래제한 대상자 ◦ 신용정보회사 신용정보 ◦ 자체 보고 자료(STR/CTR) ◦ 국내/국외 금융범죄 정보 ◦ 소셜 네트워크 서비스 ◦ 뉴스 미디어 ◦ 혐의 거래 분석 관련 자료 (이미지, 로그, 음성 녹취, 신원 확인 정보 등) ◦ 내부 고객정보 조회 보안 로그
거래 모니터링	<ul style="list-style-type: none"> ◦ 일 배치(Batch) 거래 모니터링 	<ul style="list-style-type: none"> ◦ 실시간(Real Time) 거래 모니터링
혐의 추출 모델 개발	<ul style="list-style-type: none"> ◦ 특정 기간 샘플 데이터 활용 ◦ 테스트와 검증시간 과다소요 ◦ 신속성 낮음, 정확성 모호 	<ul style="list-style-type: none"> ◦ 사용 가능한 모든 데이터 활용 ◦ 테스트와 검증 소요시간 단축 ◦ 신속성과 정확성 높음
조사 분석	<ul style="list-style-type: none"> ◦ 특정기간 거래내역 데이터를 활용한 시각화 분석 ◦ 알고 있는 혐의 유형을 찾는 ‘검색’의 문제 해결 ◦ 분석관점-속성 중심 ◦ 분석과정에서 발생하는 데이터 저장/관리 어려움 	<ul style="list-style-type: none"> ◦ 사용 가능한 모든 데이터를 활용한 빅데이터 시각화 분석 ◦ 데이터들 사이의 의미 있는 관계를 찾아가는 ‘발견’의 문제 해결 ◦ 분석관점-관계 중심 ◦ 분석과정에서 발생하는 모든 데이터 저장/관리 가능

여 거래 모니터링 모델을 설계하고 테스트하며 상급기관에 보고를 위한 목적으로 사용된다. 실제 범죄 색출에는 활용 데이터 범위의 한계로 큰 효과를 못 보고 있다. 예를 들어 거래 모니터링의 경우도 최근 채널의 다양화로 금융거래가 전자이체 등의 비대면 거래가 많아지고 지급수단이 다양화 되면서 혐의의심 거래 포착을 내부 거래 데이터만으로 실행하기에는 역부족이다.

본 연구에서 제시한 빅데이터를 활용한 자금세탁방지 시스템은 그동안 데이터 접근과 저장이 어려워 간과되고 있었던 데이터 출처들을 활용한다. 그 과정에서 하둡 인프라를 이용하여 다양한 외부 출처들로부터의 데이터를 수집, 저장하고 자금세탁 혐의거래 분석에 활용한다. 특히 잠재적 위험 거래 분석과 실시간 모니터링으로 예방적 자금세탁관리를 가능하게 한다.

아직까지 국내 뿐 만 아니라 해외에서도 빅데이터를 활용한 자금세탁방지 시스템의 사례를 찾기 힘들다. 따라서 본 연구는 기존 자금세탁방지 시스템을 분석하여 시스템이 가지고 있는 문제점을 빅데이터를 활용해 개선할 수 있는 방향을 제시한 것이 큰 의의라고 할 수 있다.

하지만 본 연구는 자금세탁방지 시스템의 구축 방법

론과 모델에 대한 연구 결과가 전무한 상황에서 자금세탁방지 시스템 구축 사례분석 연구에 그친 한계를 가지고 있었다. 특히 사례 분석 과정에서 시중 여러 금융회사의 자금세탁방지 시스템의 구축 환경을 직접 분석하고 업무 담당자의 인터뷰를 진행하는 등 보다 다양한 데이터 수집 방법을 활용했다면 더욱 정확하고 현실적인 연구가 될 수 있었을 것이다.

참 고 문 헌

- [1] 금융정보분석원, 자금세탁방지 2012 연차보고서, 2013.
- [2] 김용섭, 자금세탁방지 정보시스템의 효율적 도입방안, 삼성SDS, 2008.
- [3] 이은정, 자금세탁방지를 위한 금융회사의 과제, 투이컨설팅, 2010.
- [4] 이창주, 이진천, 자금세탁방지 정보화시스템 구축·운영경험 및 방법론, 삼일회계법인 2013.
- [5] 함유근, 채승병, 빅데이터, 경영을 바꾸다, 삼성경제

연구소, 2012.

- [6] CTPA, 세무조사관을 위한 자금세탁인지핸드북, 2009.
- [7] SAS, White Paper-은행업과 분석기술, 2012.
- [8] Computerworld, “UBS installs anti-money-laundering software”, 2002.
- [9] Gantz, J. and D. Reinsel, “Extract ing Value from Chaos”, IDC, 2011.
- [10] Khac, N., S. Markos, M. O’Neill, A. Brabazon, and M. Kechadi, “An investigation into Data Mining approaches for Anti Money Laundering,” *International Conference on Computer Engineering and Applications*, 2011.
- [11] Turner, D., M. Schroeck, and R. Shockley, *Analytics: The real-world use of big data in financial services*, IBM Institute for Business Value, 2013.
- [12] PwC, How the Financial Services Industry Can Unlock the Value in Big Data, FS Viewpoint, 2014.

저 자 소 개



김 상 완(Sang-Wan Kim)

- 2005년 : 한국방송통신대학교 (경제학 학사)
- 2014년 : 건국대학교 정보통신대학원 (경영학 석사)
- 2011년~현재 : (주)바넷정보기술 차장

· 관심분야 : 비즈니스 인텔리전스, 빅데이터 처리와 분석



함 유 근(Yu-Kun Hahm)

- 1984년 : 고려대학교 정경대학 통계학과 (경제학학사)
- 1994년 : Boston University (경영학 박사)
- 1998년~현재 : 건국대학교 경영대학 교수

· 관심분야 : 빅데이터 활용, 핀테크