정 보 보 호 학 회 지
제27권 제1호, 2017. 2

패스워드 강도 측정 방법 연구 동향

김 경 훈*, 김 승 연*, 권 태 경**

요 약

오랜 기간 널리 사용되어온 패스워드 인증 기법은 여전히 대표적인 사용자 인증 수단이지만 그 사용성과 안전성 측면에서 여러모로 부정적인 부분이 많다. 일반적으로 사용자는 기억하기 쉽도록 간단한 패스워드를 선택하는 반면 서버는 추측 공격에 대해 비교적 안전하도록 복잡한 패스워드의 사용을 권장한다. 취약한 패스워드의 사용은 전체 시스템의 안전에 큰 영향을 미치게 되므로 사용자가 패스워드를 선택하는 시점에 미리 패스워드의 강도 즉 안전성을 측정하여 피드백하기 위한 기법에 관한 연구가 다각적으로 이루어져 왔다. 또한 그 일부를 다양한 방법으로 시각화하여 이미 상용시스템에 적용하고 있다. 하지만 여전히 정확한 강도 측정과 안전한 패스워드의 사용성 제고를 위한 해결이 필요하며 따라서 이와 같은 패스워드 강도 측정 방법의 일반화를 위한 연구가 꾸준히 진행되고 있다. 본 논문에서는 텍스트 기반의 패스워드 강도 측정에 관한 연구동향을 살펴보고 분석한다.

Ⅰ. 서 론

텍스트 패스워드 인증은 가장 대표적인 사용자 인증 기법 중 하나이다. 하지만 사용자들은 텍스트 패스워드를 많이 사용함에도 불구하고 여전히 약한 강도의 패스워드를 사용하고 있다. 유출 패스워드 분석 결과 특정 구조 및 의미에 편중된 취약한 패스워드를 생성하고 있음을 알 수 있다. 웹 사이트에서는 취약한 패스워드 사용을 방지하기 위해 패스워드 생성 정책, 패스워드 미터 등을 통해 사용자가 강한 패스워드를 만들도록 유도하고 있다. 패스워드 미터는 패스워드의 강도를 피드백 해주는 지시자이나 패스워드 강도를 일관성 없이 측정하는 문제점이 존재한다[26]. 정확한 패스워드 강도 측정을 위해 다방면으로 패스워드 강도 측정에 대한 연구가 진행되고 있다. 본 연구는 패스워드 미터의 정확성과 일관성, 그리고 패스워드 미터 강도 측정 방법의 일반화에 대해 논의한다. 본 논문에서는 텍스트 패스워드 강도 측정 방법 분류를 통해 연구 동향을 알아본다.

Ⅱ. 연구 배경

텍스트 패스워드는 과거부터 현재까지 가장 많이 사용되고 있는 사용자 인증 기법이다. 기억에 의존하는 경향이 강하므로, 사용자들은 대체로 기억하기 쉬운 패스워드를 생성하고 있다. 사용자는 패스워드를 기억하기 쉽도록 키보드 배열 패턴, 자신과 관련한 단어, 숫자 등 유추하기 쉬운 패스워드를 사용하고 이를 재사용하고 있다[1, 2, 3, 4, 5]. 즉, 기억하기 쉬운 패스워드는 대체로 취약한 패스워드이다. 사용자가 이러한 취약한 패스워드를 선택하는 것을 방지하기 위해 서비스 제공자는 복잡한 패스워드, 다시 말해 강한 패스워드 생성을 유도하고 있다. 강한 패스워드를 생성하도록 하는 방법 중 하나로 시스템이 패스워드를 생성해주는 방법이 있으나, 이는 사용자들이 기억하기 어렵다는 문제점이 있다. 그러므로 사용자가 직접, 그리고 강한 패스워드를 선택하도록 유도하기 위해 웹 사이트와 같은 서비스 제공자는 패스워드 생성 정책 및 패스워드 미터를 제공하고 있다. 패스워드 미터는 먼저 패스워드의 강도를 측정한 후 그 결과를 다양한 지시자를 통해 사용자에게 피드백 해주는 기능을 의미한다[28]. [표. 1]과 같이 패스워드 미터는 막대그래프, 문자, 그림 지시자를 이용하여 패스워드 강도를 피드백한다. [표 1]에서 볼 수 있듯이 서로 다른 표현과 단어를 이용하고 있고, 또한 강도 스케일이 통일되지 않고 있는 문제점이 있다.

본 연구는 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입니다. (No. NRF-2015R1A2A2A01004792)
* 연세대학교 정보대학원 정보보호연구실 ({rickyboss, tribunus000}@yonsei.ac.kr)
** 교신저자, 연세대학교 정보대학원 정보보호연구실 (taekyoung@yonsei.ac.kr)

〔표 1〕 패스워드 미터 종류

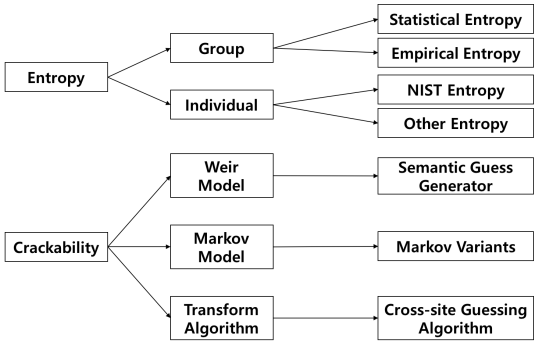
웹 사이트	패스워드 미터 표현 방법 (지시자)
Google	비밀번호 안전성: 너무 짧음 비밀번호 안전성: 안전하지 않음 비밀번호 안전성: 적합함 비밀번호 안전성: 강력
Microsoft	사용 불가능 / 사용 가능
Naver	사용불가 적정 낮음 높음
Nate	사용불가 보안상태안전 비밀번호는 6자리 이상으로 입력해 주세요. 사용가능한 안전한 비밀번호 입니다. 보안상태낮음 보안상태높음 좀 더 안전한 비밀번호로 설정해 주세요. 예측이 어려운 강력한 비밀번호 입니다
Apple	안전 정도: 안전 정도: 보통임 안전 정도: 안전함

기존 연구[26]에 따르면 시각적인 면에서 패스워드 미터의 일관되지 않은 표현 문제와 더불어, 강도를 일관성 없이 측정하고 있는 문제가 존재한다. 해당 연구에서는 유출 패스워드(Rockyou, Phpbb, Top 500), 패스워드 사전을 이용하여 패스워드 미터의 비일관성 문제를 밝혔다. 동일한 패스워드를 입력함에도 불구하고, 서로 다른 웹 사이트에서 서로 다른 패스워드 강도를 피드백하고 있음을 밝혔다. 국내 웹 사이트 패스워드 미터 연구에서도 패스워드 미터가 일관성 없는 강도 측정을 하고 있음을 알 수 있다[29]. 서비스 제공자는 다양한 방법으로 강한 패스워드를 생성하도록 유도하고 있지만 패스워드 강도 측정의 기준이 마련되어 있지 않다는 한계점이 존재한다.

강한 패스워드 생성을 위해 텍스트 패스워드 강도 측정과 관련한 연구들이 이루어지고 있다. 본 논문에서는 텍스트 패스워드 강도 측정 방법 분류를 통해 연구 동향에 대하여 분석한다.

Ⅲ. 패스워드 강도 측정법 분류

과거에는 전수공격을 통해 패스워드를 추측하였으며, 길이, 문자 종류에 초점을 맞춰 패스워드 엔트로피를 결정하였다. 그러나 사용자들이 기억하기 쉬운 단어, 패턴을 사용함에 따라 엔트로피는 높지만 실제로는 취약한 패스워드가 생겨나게 되었다. 단순한 방식으로 패스워드 강도를 측정하기 어려워짐에 따라 새로운 강도 측정 방법이 나타나게 되었다.



〔그림. 1〕 엔트로피와 Crackability의 구분

본 연구에서는 패스워드 강도 측정 방법을 엔트로피와 crackability로 나누어 분류한다([그림. 1]). 엔트로피는 사용자가 입력한 패스워드가 알고리즘을 통과했을 때 나오는 정보량의 단위가 bit인 경우를 일컫는다. 엔트로피는 검증 대상에 따라 그룹 엔트로피와 개별 엔트로피로 구분한다. crackability는 입력한 패스워드가 알고리즘을 통과했을 때 출력되는 결과가 크래킹 될 가능성인 경우를 일컫는다. 알고리즘 사용 모델에 따라 PCFG 모델, 마코프 모델, 변환 알고리즘으로 구분한다.

3.1. 패스워드 엔트로피

본 연구에서의 엔트로피는 '어떤 패스워드가 알고리즘을 통과했을 때 나오는 정보량의 단위가 비트'인 경우를 의미하며, 크게 개별 엔트로피, 그룹 엔트로피로 구분한다. 개별 엔트로피는 패스워드 각각에 대한 엔트로피를 구할 수 있는 알고리즘을 의미한다. 그룹 엔트로피는 한 패스워드 집단 전체에 대하여, 그 집단의 분포를 통해 엔트로피를 구하는 방법이다. 서로 다른 특성을 가진 패스워드 그룹 간 비교가 용이하다는 장점이 있지만, 집단의 크기가 커야하므로 표본이 많이 필요한 단점을 가지고 있다.

3.1.1. Shannon 엔트로피와 변형 엔트로피

엔트로피는 열역학에서 에너지의 흐름을 설명할 때 이용되는 상태 함수이며, 어떤 확률 분포에 담긴 불확실성의 한 척도이다. 정보이론에서의 엔트로피는 1948년

Shannon이 처음 도입하여[6], 정보 엔트로피 또는 샤논 엔트로피라 불렸으며, 이를 기반으로 다양한 연구들이 나오게 되었다. 샤논 엔트로피는 하나의 패스워드가 가지고 있는 정보량의 비트 수를 나타내며, 불확실성이 높아질수록 정보의 양은 더 많아지고, 엔트로피 값은 증가한다.

$$H(X) = -\sum_{x} P(X=x)\log_2 P(X=x) \quad (1)$$

샤논 엔트로피를 시작으로 많은 변형이 나오게 되었고, 그 중 샤논 엔트로피를 일반화 한 Renyi 엔트로피가 있다[7]. 샤논 엔트로피는 Renyi 엔트로피의 n → 1 한 결과이다.

$$H_n(X) = \frac{1}{1-n} lg(\sum_{i=1}^{N} P_i^n) \quad (2)$$

Renyi 엔트로피 또한 많은 변형이 있으며, 다음과 같은 종류가 있다.

- Hartley 엔트로피[8] : Renyi 엔트로피의 식에 n = 0 을 대입한 것으로 샤논 엔트로피보다 이전에 연구되었으며, 분포의 사이즈만 고려한 엔트로피 이다.
- Collision 엔트로피 : Renyi 엔트로피의 식에 n = 2 인 결과를 나타낸다.
- Min 엔트로피 : Renyi 엔트로피 식에서 n을 무한대로 보낼 때 나오는 값의 분포에서 가장 많이 나오는 확률에 기반을 두는 엔트로피 계산이다.

3.1.2. 그룹 엔트로피

과거에는 적은 수의 패스워드로 전체 확률 분포를 가정하여 패스워드 강도를 구하기 위한 통계학적 근거를 사용하였다. 초기에는 패스워드 정책이 복잡하지 않았기 때문에 문자 종류, 패스워드 길이에 근거한 단순한 방법으로 엔트로피를 계산하였지만, 정책이 복잡해짐에 따라 다양한 통계학적 방법으로 엔트로피를 계산하게 되었다. 그룹 엔트로피는 분포를 가정하는 방법에 따라 통계학적 엔트로피와 경험적 엔트로피로 구분된다.

- 통계학적 엔트로피 : 패스워드의 일부를 이용하여 전체의 패스워드 분포를 추측하여 엔트로피를 구하는 방식이다. 극단적인 이상점은 무시되고, 전체 분포의 경향성을 따라가게 된다. 초기에는 균등 분포로 가정하였으나 점차 심도 있는 통계학적인 방법을 이용하여 엔트로피를 구하는 방법으로 연구되고 있다. 그룹 엔트로피를 구하면 다른 특성을 가진 패스워드 집단끼리 비교하기 용이한 장점을 가지고 있다.
- 경험적 엔트로피 : 샤논 엔트로피 기반 엔트로피로, 전체 집합 중 특정 단어가 나올 확률을 계산하여 구하는 방법이다. 2010년 Shay 등의 연구에서 처음 사용되었다[9].

통계학적 엔트로피는 Florencio 등의 첫 연구 결과가 있다[4]. 툴바를 이용하여 패스워드를 수집하였고, 패스워드 엔트로피로 분류하여 패스워드의 강도를 비교하였다. 균등 분포를 가정하여 계산식 $\log_2((char.size)^{len})$을 이용하여 엔트로피를 계산하였다. 연구 결과 New York Times 웹 사이트에서 사용하는 패스워드의 경우 37.68 비트에 비해 MS Outlook Web Access의 엔트로피는 51.36비트로 사이트 별 패스워드 강도의 차이가 있음을 밝혔다. 이후, Egelman 등의 연구에서는 이전 Florencio 등의 연구와 비슷한 계산법을 사용하여 강도를 측정하였고, zero-order 엔트로피라 이름을 붙였다[10]. 알파벳 빈도가 고려되지 않는 한계점이 존재하지만 현재의 패스워드 미터가 기반하고 있는 방법이며, 정책에 따른 패스워드의 상대적인 차이를 측정하기에 적합한 방법이다. 통계학적인 방법으로 패스워드 그룹의 분포를 추측하여 정확성을 높인 방법이 존재한다. Bonneau의 연구에서는 포아송 분포 등 다양한 통계 분포와 부분 추측 메트릭을 이용하여 더 적은 추측 횟수로 효율성을 높일 수 있는 방법을 제안하였다[11]. 또한 Malone 등은 패스워드 분포를 Zipf 분포로 모델링한 연구를 진행하였다[12]. Zipf 분포는 패스워드의 선택과 자연어 단어 사용에 상당한 차이가 나는 상황에서 패스워드 분포의 기울어진 현상을 설명할 수 있는 가능성이 있는 분포이다. Zipf 분포로 모델링을 진행 한 후, 추측작업, 샤논 엔트로피, Renyi 엔트로피, 최소 엔트로피 등을 사용하여 실제 분포와 모델링 방법 별로 엔트로피가 어떻게 차이나는지 분석하였다. 그 결과, Zipf 분포로 모델링을 한 것과 실제 데이터 사이의 비교 결과에서 더 높은 효율을 보이고 있음을 밝혔다.

경험적 엔트로피는 Shay 등의 연구에서 처음 사용되었다[9]. 카네기멜론 대학교 구성원 중 109명의 설문 응답을 통해, 길이, 숫자의 개수, 특수문자의 위치, 사전 단어 사용 여부 등을 파악하여 사용하는 패스워드의 엔트로피를 계산하였다. 만약 참여자 중 28%의 사람들이 패스워드 길이 8글자를 가지고 있다고 대답하면, 어떠한 패스워드의 길이가 8일 확률은 0.28이 되고, 이를 각각의 길이에 대하여 구한 후, 샤논의 엔트로피 공식에 넣어 더하는 방식으로 엔트로피를 구한다. 확률 계산 시 대소문자, 숫자, 특수문자의 개수, 위치, 사전 의미 등에 따른 확률을 계산하였으며, 사전 단어에 관련한 사항은 NIST 엔트로피에서 사용하는 방식을 참고하였다. 최종 결과는 집합 전체의 엔트로피이다. Shay 이후, 2011년 Komanduri 등의 연구에서도 같은 방법으로 그룹 별 엔트로피를 계산하는 연구를 진행하였다[13]. 서로 다른 패스워드 정책에 따라 경험적 엔트로피가 어떠한 차이가 있는지 분석하였다. 그 결과, 복잡하거나 긴 패스워드를 요구하는 정책의 엔트로피가 높음을 밝혔다.

3.1.3. 개별 엔트로피

개별 엔트로피는 패스워드 각각의 엔트로피를 비트로 나타내는 방식이다. 통계학적인 분포를 가정하기에는 어려움이 있지만, 개별 엔트로피를 구하려는 다양한 시도가 있다. 특히 샤논 엔트로피에 기반을 둔 NIST 엔트로피는 단순한 계산으로 각각의 패스워드 엔트로피를 구할 수 있어 여러 연구에서 널리 사용되고 있다. 하지만 패스워드의 다양성을 고려 할 수 없다는 한계점이 있고 가정된 확률분포에 근거하기 때문에 몇몇 연구에서는 NIST 엔트로피가 과대평가된 결과를 주는 한계점을 가지고 있다[15, 16].

NIST 엔트로피는 미국국립표준기술연구소에서 제안한 엔트로피로 SP800-63 문서에 언급되어 있다[14]. 패스워드의 길이와 각 자리에 사용 가능한 문자수에 기반을 둔 엔트로피이다. 패스워드 분포도를 가지고 있는 공격자가 상대방의 패스워드를 추측하기 위해 가장 사용 가능성 높은 패스워드부터 시도하여 확률을 줄여나갈 때 평균 양을 의미한다. NIST 엔트로피를 구하는 방법은 다음과 같다.

1). 첫 문자의 엔트로피는 4.6비트, 두 번째부터 여덟 번째 자리 문자까지 엔트로피는 2.3비트, 아홉 번째부터 스무 번째 자리까지는 1.5비트, 스물한 번째 자리 이상은 엔트로피 1비트를 각각 부여한다.

2). 쉬운 단어부터 시도하는 공격자의 공격을 막기 위하여 구성 검사와 사전 검사를 거친 패스워드에 각각 6비트의 엔트로피를 추가한다. 구성 검사는 패스워드 내에 대문자, 소문자, 특수문자, 숫자를 섞은 패스워드를 만들기 위함이고, 사전 검사는 자주 사용되는 영어 단어 사용을 금지하기 위함이다. 16자 이상의 패스워드는 사전검사를 적용하지 않는 한계점이 있다.

3). 길이 검사결과의 비트와 구성 검사, 사전 검사 결과의 비트를 모두 더한 값이 NIST 엔트로피의 최종 비트값이 된다.

NIST 엔트로피는 길이와 문자열의 개수에만 의존하기 때문에 사전이나 블랙리스트에 나오는 약한 패스워드를 인지 할 수 없는 한계점이 있다. 이에 2010년 Weir 등의 연구에서는 정확한 샤논 엔트로피 값이라 할지라도, 온라인 패스워드 크래킹 공격에 얼마나 취약한지 알 수 없기에 crackability를 제안하였다[15].

3.2. 패스워드 crackability

Crackability는 '어떠한 패스워드가 크래킹 될 가능성'로 정의한다. Crackability는 제한된 자원과 유출된 패스워드를 이용하여 훈련한 결과를 이용하기 때문에 더 현실적인 결과를 보여주는 장점이 있다. 대부분 연구에서는 어떠한 알고리즘에 패스워드 입력 시, 추측횟수에 따라 크래킹 될 가능성에 대하여 연구한다. 알고리즘은 크게 PCFG 기반 알고리즘과 마코프 체인 (Markov chain) 알고리즘으로 구분된다. 두 알고리즘 모두 자연어 처리에 사용되는 알고리즘이며, 패스워드가 일상생활의 언어와 다르지 않다는 생각에 기반을 두어 사용되고 있다.

3.2.1. PCFG (Probabilistic Context Free Grammar)

PCFG 는 2009년 Weir 등에 의해 처음 패스워드 연구에 적용되었다[17]. 패스워드의 구조를 파악한 후, 각 자리에 특정 문자가 올 확률을 계산하는 방식이다. 맹글링 (Mangling) 규칙을 익히고 훈련 데이터를 기준으로 확률을 도출 한 후, 최적의 확률 순으로 추측을 한다. Weir 모델의 특징은 길이가 정확히 맞아 떨어지는 단어

와의 정합을 찾는 것이다. 만약 구조가 'LLLLDLLLS(L: 소문자, D: 숫자, S: 특수문자)'라면 처음의 'LLLL'은 have, cats, dogs 등과 같이 딱 4자리인 것을 찾는 것이다.

Weir의 알고리즘을 이용하여 Kelley의 연구에서는 패스워드 구조, 문자, 숫자, 특수문자의 패턴에 따른 확률에 따라 추측 순서를 결정하는 연구를 진행하였다[16]. 또한 구조에 맞는 하위 문자열의 확률 순으로 더욱 정교한 추측 순서를 결정한다. 터미널은 구체적인 하위 문자열을 갖는 구조이다. 훈련 데이터로 검색 테이블(lookup table)을 만들고 각 패스워드의 추측 횟수를 계산한다. 검색 테이블 생성은 추측 횟수가 50조를 넘어가면 무시한다. 훈련 데이터로 패스워드 구조와 터미널을 익혀 패스워드가 추측될지, 추측된다면 몇 회의 추측이 필요할지 예상하여 계산한다. Ur의 연구에서는 변형된 4가지 PCFG 버전을 이용하여 패스워드 추측에 대한 연구를 진행하였다[18]. 연구 결과에서는 PCFG 중 테스트 할 패스워드의 반을 훈련 데이터로 사용하는 것이 전반적으로 crackability가 높음을 밝혔다. 시맨틱 토큰과 POS (parts-of- speech) 태깅을 이용하여 문법적인 구조를 모델링하였다[19]. POS 태깅은 각 단어가 명사, 동사, 형용사 등 어떤 품사인지 정하는 과정을 말하며, CLAWS를 이용하여 각 단어에 태그를 할 수 있다. 긴 패스워드의 문법적인 구조를 이용하여 검색 량을 감소와 그 안의 많은 단어를 자동적으로 혼합하여 크래킹하는 기법을 소개하였다. 긴 패스워드를 효율적으로 크래킹 할 수 있으며, 최신 패스워드 크래커가 크래킹하지 못한 전체 집합의 10% 정도 되는 패스워드를 추가적으로 크래킹 하는 성능을 보였다. PCFG에 기반을 둔 의미 구에 대한 연구는 2014년 Veras 등의 연구에서 본격화 되었다[2]. 자연어 처리 방법을 이용하여 패스워드의 의미 구를 나누고 분류하며, COCA (Contemporary Corpus of American English)와 영어 이름, 도시, 성, 나라이름 등을 가지고 훈련을 시킨다. 구문과 시멘틱 카테고리를 분류하여 문법의 구문론에 따라 분류로 나타내었다. 기존의 패스워드 크래킹 방법에 비해 유출 패스워드 LinkedIn에서 67%, MySpace 32%를 더 크래킹하는 성능을 보였다.

3.2.2. 마코프 체인 (Markov Chain) 모델

마코프 체인은 자연어 처리 방법 중 하나인 알고리즘이며, 한 단어를 위수에 따라 n-그램하여 어떤 단어가 나올 확률을 조건부 확률로 예측하는 방법이다. n-그램 마코프 모델은 하나의 문자 안에서 길이 n인 접두사에 기반을 두어 다음에 올 문자에 대한 확률을 모델링하는 것이다. 즉, 패스워드 전체에 대해 n-그램 방식을 적용한다.

마코프 체인 모델을 패스워드 크래킹에 도입한 것은 '기억하기 쉬운 패스워드는 자연어의 특성과 비슷할 것'이라는 아이디어에서 시작되었으며, 2005년 Narayanan이 최초로 패스워드 크래킹에 도입하였다[20]. 마코프 체인 모델의 다양한 조건 중 위수-1(바로 이전 문자 하나에 의존)하는 방안이 많이 연구 되고 있다.

Narayanan은 패스워드 검색 공간의 사이즈를 줄여 효율적으로 크래킹 할 수 있는 환경을 구축하였다[20]. 문자로 이루어진 세그먼트의 확률을 구하기 위하여 패스워드의 구조를 파악하고 각 세그먼트 별 n-그램을 사용하였다. Narayanan 연구 역시 위수-1 외의 다른 조건은 배제하였다. 이후, 2010년 Dell 등의 연구에서는 Narayanan의 연구 방식을 사용하여, 유출 패스워드를 분석하였다. 마코프 체인 모델 기반의 사전 공격, 사전 기반 맹글링, PCFG 맹글링 등 다양한 크래킹 알고리즘을 이용하여 비교 연구를 진행하였다[21]. 검색 공간의 사이즈에 따른 차이는 원본의 데이터 셋과 큰 차이가 존재하지 않다는 것을 밝혔다. 사전 공격 혹은 맹글링 기술은 보다 적은 공간에 있어서 더 나은 결과를 도출함을 밝혔다.

Castelluccia 등의 연구에서는 대용량 사전 데이터를 이용하여 whole string 마코프 모델을 이용하는 방안을 제안하였다[22]. 제안한 강도 측정 방안은 기존의 패스워드 측정 방법에 비해 더 높은 정확성을 가지고 있음을 밝혔다. 최근의 마코프 체인에 관한 연구에서는 위수 조건과 smoothing 기법을 적용한 연구가 진행되었다[23]. 각 기법에 대하여 Rockyou, Phpbb, Yahoo 등의 데이터를 이용하여, Weir 알고리즘, JtR(John the Ripper)와 비교를 통해 마코프 모델의 성능을 검증하였고, 마코프 모델의 성능이 다른 모델에 비해 더 나은 성능을 보이고 있음을 밝혔다. 그 후, Durmuth 등의 연구에서는 마코프 모델에 smoothing 기술을 추가한 연구가 진행되었고[24], 마코프 변형 연구가 최근까지 이어

지고 있다[1].

3.2.3. 변환 알고리즘

변환 알고리즘은 '이전에 사용했던 패스워드를 알게 될 경우 다음의 패스워드를 추측할 수 있을 것'이라는 아이디어에서 시작된다. Zhang 등의 연구에서는 과거의 패스워드를 알고 있을 경우 다음에 사용될 패스워드를 추측하는 알고리즘을 제안하였다[25]. 실제 학교를 대상으로 SSO 시스템의 과거 패스워드를 수집하여 실험하였다. 알고리즘에서는 과거 패스워드의 변경 가능한 방법을 트리로 표현하여, 깊이에 따라 약 2가지의 변형 가능한 가지 수를 보였다. 연구 결과에서는 리트 변환(ex. s → $)과 키 변환(ex. 1 → !)을 혼합하여 추측하는 것보다 과거에 사용한 변형방법과 혼합하여 추측하는 것이 더 좋은 결과를 낼 수 있다는 것을 밝혔다. Das 등의 연구에서는 한 사람이 여러 사이트를 이용할 때 사이트마다 비밀번호가 어떻게 연관되어 있는지 연구하였다[3]. 유출 패스워드 중 같은 식별자를 가지는 것끼리 묶은 후 패스워드의 유사성을 분석하였다. 그 결과 한 사용자가 여러 사이트를 이용할 때 패스워드를 재사용하거나 수정하는 것을 밝혀냈다. 이를 기반으로 하여 입력 값을 넣으면 입력 값과 비슷하지만 다르게 변형해 주는 cross-site 패스워드 추측 알고리즘을 제안하였다.

Ⅳ. 최근 패스워드 강도 측정 연구

Dropbox Inc.의 Daniel은 2016년 경량형 패스워드 미터인 zxcvbn[27]을 발표했다. zxcvbn은 일반적으로 널리 쓰이는 패스워드 조건인 3class8과 NIST가 제안한 강도 측정 방식의 문제점을 보완함과 동시에 정확한 패스워드 강도를 진단 하고자 하였다. zxcvbn은 확률에 의해 추측가능성을 추정하지 않으며, 경험적으로 추정한다. 패스워드가 주어지면 매칭, 측정, 검색 단계를 거쳐 패스워드 추측 시 필요한 추측 시도 횟수를 진단하며, 가정하는 공격자는 패스워드에 특정 패턴이 포함되었는지 여부를 알고 있는 공격자이다. zxcvbn은 위의 과정을 거쳐 입력된 패스워드의 크래킹 예상 시간을 산출하고, 그 결과를 세기, 연, 월, 일, 시간, 분, 초 단위로 출력한다.

Ⅴ. 결 론

기존 연구들은 패스워드 미터를 이용한 강도 측정 과정이 일관적이지 않은 문제를 다양한 관점에서 밝혔다. 본 논문에서는 이러한 기존 연구를 패스워드 강도 측정 방법에 따라 크게 엔트로피, crackability 관점으로 나누어 연구 동향을 분석하였다.

추후 연구로는 다양한 웹 사이트의 패스워드 미터의 판정 결과가 실제로 정확한가에 대해 유출 패스워드와 crackaability 알고리즘, 크래킹 툴 등을 활용한 실증적 검증이 가능하다. 또한 그 결과로부터 밝혀진 문제점을 보완함으로써 보다 정확하며 일관적인 강도 측정 방법을 도출해낼 수 있을 것이다.

참 고 문 헌

[1] B. Ur, F. Noma, J. Bees, S. M. Segreti, R. Shay, L. Bauer, N. Christin, and L. F. Cranor, ""I added '!' at the End to Make It Secure":Observing Password Creation in the Lab," in Proc. of SOUPS, 2015.

[2] R. Veras, C. Collins, and J. Thorpe, "On the Semantic Patterns of Passwords and their Security Impact," in Proc. of NDSS, 2014.

[3] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. F. Wang, "The Tangled Web of Password Reuse," In Proc. of NDSS, 2014.

[4] D. Florencio and C. Herley, "A Large-Scale Study of Web Password Habits," in Proc. of WWW, 2007.

[5] S. Gaw and E. W. Felten, "Password Management Strategies for Online Accounts," in Proc. of SOUPS, 2006.

[6] C. E. Shannon, "A mathematical theory of communication," ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), pp. 3-55, 2001.

[7] A. Rrnyi, RNYI, "On measures of entropy and information," In: Fourth Berkeley symposium on mathematical statistics and probability, pp. 547-561, 1961.

[8] R. V. Hartley, "Transmission of information1,"

Bell System technical journal, 7(3), pp. 535-563, 1928.

[9] R. Shay, S. Komanduri, P. G. Kelley, P. G. Leon, M. L. Mazurek, L. Bauer and L. F. Cranor, “Encountering stronger password requirements: user attitudes and behaviors,” In Proceedings of the Sixth Symposium on Usable Privacy and Security ACM. July, 2010.

[10] S. Egelman, A. Sotirakopoulos, I. Muslukhov, K. Beznosov, and C. Herley, “Does my password go up to eleven?: the impact of password meters on password selection,” In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ACM, pp. 2379-2388. April, 2013.

[11] J. Bonneau, “The science of guessing: analyzing an anonymized corpus of 70 million passwords,” In 2012 IEEE Symposium on Security and Privacy, pp. 538-552, May, 2012.

[12] D. Malone, and K. Maher, “Investigating the distribution of password choices,” In Proceedings of the 21st international conference on World Wide Web, ACM, pp. 301-310, April, 2012.

[13] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, and S. Egelman, “Of passwords and people: measuring the effect of password-composition policies,” In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 2595-2604, May, 2011.

[14] W. E. Burr, D. F. Dodson, E. M. Newton, R. A. Perlner, W. T. Polk, S. Gupta, and E. A. Nabbus, “Sp 800-63-1. electronic authentication guideline,” NIST, 2013.

[15] M. Weir, S. Aggarwal, M. Collins, and H. Stern, “Testing metrics for password creation policies by attacking large sets of revealed passwords,” In Proceedings of the 17th ACM conference on Computer and communications security, pp. 162-175, October, 2010.

[16] P. G. Kelley, S. Komanduri, M. L. Mazurek, R. Shay, T. Vidas, L. Bauer, and J. Lopez, “Guess again (and again and again): Measuring password strength by simulating password-cracking algorithms,” In 2012 IEEE Symposium on Security and Privacy, pp. 523-537, May, 2012.

[17] M. Weir, S. Aggarwal, B. De Medeiros, B., and B. Glodek, “Password cracking using probabilistic context-free grammars,” In 2009 30th IEEE Symposium on Security and Privacy, pp. 391-405, May, 2009.

[18] B. Ur, S. M. Segreti, L. Bauer, N. Christin, L. F. Cranor, S. Komanduri, R. Shay, “Measuring real-world accuracies and biases in modeling password guessability, "In 24th USENIX Security Symposium, pp. 463-481, 2015.

[19] A. Rao, B. Jha, and G. Kini, “Effect of grammar on security of long passwords,” In Proceedings of the third ACM conference on Data and application security and privacy, pp. 317-324, February, 2013.

[20] A. Narayanan, and V. Shmatikov, “Fast dictionary attacks on passwords using time-space tradeoff,” In Proceedings of the 12th ACM conference on Computer and communications security, pp. 364-372, Nov. 2005.

[21] Dell' Amico, Matteo, P. Michiardi, and Y. Roudier, “Password Strength: An Empirical Analysis,” In INFOCOM, Vol. 10, pp. 983-991, March, 2010.

[22] C. Castelluccia, M. Dürmuth, and D. Perito, “Adaptive Password-Strength Meters from Markov Models,” In NDSS, Feb., 2012.

[23] J. Ma, W. Yang, M. Luo, and N. Li, “A study of probabilistic password models,” In 2014 IEEE Symposium on Security and Privacy, pp. 689-704, May, 2014.

[24] M. Dürmuth, F. Angelstorf, C. Castelluccia, D. Perito, and A. Chaabane, “OMEN: Faster password guessing using an ordered markov enumerator,” In International Symposium on

Engineering Secure Software and Systems, pp. 119-132, March, 2015.

[25] Y. Zhang, F. Monrose, and M. K. Reiter, “The security of modern password expiration: An algorithmic framework and empirical analysis,” In Proceedings of the 17th ACM conference on Computer and communications security, pp. 176-186, Oct., 2010.

[26] X. de C. de Carnavalet and M. Mannan, “From Very Weak to Very Strong: Analyzing Password-Strength Meters,” In Proc. of NDSS, 2014.

[27] D. L. Wheeler, “zxcvbn: Lowbudget password strength estimation,” In Proc. of 25th USENIX Security Symposium, pp. 157-173, 2016.

[28] B. Ur, P. G. Kelley, S. Komanduri, J. Lee, M. Maass, M. L. Mazurek, and L. Christin, “How does your password measure up? The effect of strength meters on password creation,” In USENIX Security Symposium, pp. 65- 80, Aug., 2012.

[29] 김경훈, 권태경, “김경훈, 권태경, “국내 웹 사이트 패스워드 미터 분석,” 정보보호학회논문지, Vol. 26, No. 3, pp. 757-767, 2016.

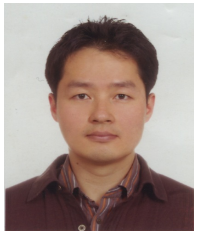
〈저 자 소 개〉

김 경 훈 (KyoungHoon Kim)
학생회원

2015년 2월 : 성공회대학교 컴퓨터공학 학사
2015년 3월~현재 : 연세대학교 정보대학원 석사과정
관심분야 : Authentication, Usable Security 등

김 승 연 (Seung-Yeon Kim)
학생회원

2015년 2월 : 세종대학교 응용통계학 및 컴퓨터공학 학사 (자연과학대학 수석졸업)
2015년 3월~현재 : 연세대학교 정보대학원 석박통합과정
관심분야 : Usable Security, Social Engineering, 스마트폰 보안



권 태 경 (Taekyoung Kwon)
종신회원

1992년 2월 : 연세대학교 컴퓨터과학과 학사
1995년 2월 : 연세대학교 컴퓨터과학과 석사
1999년 8월 : 연세대학교 컴퓨터과학과 박사
1999년~2000년 : U.C. Berkely Post-Doc.
2001년~2013년 8월 : 세종대학교 컴퓨터공학과 교수
2007년~2008년 : Univ. Maryland at College Park 교환교수
2013년 9월~현재 : 연세대학교 정보대학원 교수
관심분야 : 암호프로토콜, 네트워크 프로토콜, 사물인터넷 보안, Usable Security, HCI 등