

룰 기반 CRM 시스템에서 시간요소를 고려한 룰 모델 제안

황혜수⁰, 이태희, 이상구
서울대학교 컴퓨터공학부
{hstg, thlee, sglee} @europa.snu.ac.kr

A Time Component-Aware Rule Model For Rule-Based Customer Relationship Management Systems Heasoo Hwang⁰, Taehee Lee, Sang-goo Lee School of Computer Science and Engineering, Seoul National University

요 약

최근 고객에게 개인화된 서비스를 제공하기 위한 방법의 하나인 CRM 시스템에 대한 연구가 활발히 이루어지고 있다. 본 논문에서는 룰 기반 CRM에서의 효율적인 룰 매칭을 목적으로 하는 시간요소를 고려한 룰 모델을 제안하고자 한다. 룰 매칭 시점관련조건을 고객관련조건과 구분함으로써 룰의 의미를 구조적으로 명확히 파악할 수 있고 룰 매칭 시에 룰 매칭 시점관련조건을 고객관련조건보다 먼저 수행할 수 있으므로 현재 시점에 룰 매칭을 해야 하는 룰들만을 대상으로 전체 고객 데이터를 검색할 수 있다. 이 때 룰 매칭 시간요소조건을 분류하기 위하여 필요한 카테고리 정보를 저장하고 계산 로직의 반복 수행을 막기 위한 목적으로 어휘를 정의하여 사용하는 방법에 대해서 설명한다.

1. 서론

최근 몇 년간 고객에게 개인화된 서비스를 제공하기 위한 여러 방법들[1,2,3]이 급속히 발전하여 왔다. 특히 CRM(Customer Relationship Management)과 관련해서는 많은 연구들이 활발히 진행되고 있다[4].

CRM이란 고객에 대한 정확한 이해를 바탕으로 고객이 원하는 제품과 서비스를 적절한 시기에 제공하는 통합 고객 관리이다. CRM의 궁극적인 목적은 "(1)The right offer (2)To the right person (3)At the right time (4)Through the right channel" 이라 정의할 수 있다[5]. CRM 룰에 나타나는 조건들은 위 네 가지 사항을 표현하기 위한 것들이므로 위의 (1)~(4)를 CRM 조건들의 카테고리라 정의하도록 하겠다. 룰 기반 CRM 시스템에서는 네 카테고리의 조건들을 하나의 룰로 표현하고 이렇게 정의한 룰에 따라 고객 대상 캠페인이 자동적으로 수행될 수 있어야 한다. 따라서 CRM을 위한 룰의 구조는 CRM 조건들을 효과적으로 표현할 수 있어야 하고, 동시에 고객 데이터와 룰 간의 매칭 시간을 단축시킬 수 있는 것이어야 한다.

그러나 현재의 룰 기반 CRM 시스템들은 룰의 조건이 위 네 카테고리 중 어디에 속하는지를 명확히 구분하지 않고 있다. 따라서 실제로는 카테고리가 다른 조건들이 하나의 IF절에 섞여 있게 됨으로써 사람이 룰의 의미를 정확하게 파악하기 힘들다. 특히 룰 매칭 시점관련 조건이 고객관련조건으로부터 명확히 분리되지 않았기 때문에 룰 매칭 시에 룰 매칭 시점에 관한 고려 없이 모든 룰들을 매칭하게 된다. 따라서 본 논문에서는 시간관련조건을 고객관련조건과 분리하여 저장하는 룰 구조를 제시하고 조건의 카테고리 구분을 위하여 어휘를 정의하여 조건에 사용하도록 하였다.

본 논문의 구성은 다음과 같다.

2.1절에서는 CRM에서의 룰 매칭 시스템을 간략히 소개한다. 그리고 2.2절에서는 룰의 의미를 구조적으로 파악할 수 있고 매칭할 룰의 개수를 줄이기 위하여 룰 수행 시점에 관련된 시간요소조건을 고객관련조건과 분리하는 방안을 제시했다. 룰 수행 시점에 관련된 시간요소 조건이란 위 네 가지 CRM의 목표 중에서 "at the right time"을 달성하기 위한 조건에 해당한다. 룰의 시간요소조건을 분리하게 되면 이를 만족되는 룰에 대해서만 고객 데이터 검색을 해도 되므로 고객 데이터와 매칭해야 할 룰의 개수가 줄어들게 되므로 전체 룰 매칭 시간이 단축된다.

위와 같이 시간요소조건을 분리하기 위해서는 그 조건의 카테고리 룰 알 수 있어야 한다. 조건의 카테고리는 조건에 사용하는 용어의 카테고리에 해당한다. 그러나 조건에 데이터베이스 필드를 사용하게 되면 그 필드의 카테고리 관련 정보는 존재하지 않으므로 알 수 없다. 따라서 2.3절에서는 CRM 조건에 사용하는 용어로서 데이터베이스의 기본 데이터 필드 대신 용어 카테고리 정보를 담은 어휘를 정의하도록 한다. 어휘 메타 데이터에는 카테고리 정보 뿐만 아니라 어휘 데이터 값을 실제로 생성해 내는 계산 로직 정보도 포함시킴으로써 매칭 속도를 향상시킬 수 있다.

2. 시간요소를 고려한 룰 모델

2.1 전체 CRM 룰 시스템의 구성

CRM 룰 매칭 엔진은 룰과 고객 데이터를 비교하여 룰 조건을 만족하는 고객들의 리스트를 찾아내는 역할을 한다. 캠페인 관리 시스템에서는 룰 매칭 엔진에서 생성한 고객 리스트를 기반으로 그 룰에 해당하는 캠페인을 수행하게 된다.

그림 1의 룰 매칭 엔진의 구조는 다음과 같다.

- ① 룰의 유효성을 검사하는 모듈
- ② 룰의 품질을 예측하는 모듈
- ③ 룰의 시간요소조건을 검사하는 모듈
- ④ 룰의 고객관련조건을 검사하는 모듈
- ⑤ SQL문의 조인조건을 생성하는 모듈
- ⑥ 패턴 룰(association/sequential 룰)을 매칭하는 모듈
- ⑦ 캠페인 관리 시스템과의 연결 모듈
- ⑧ 룰 에디터에 제공하는 API

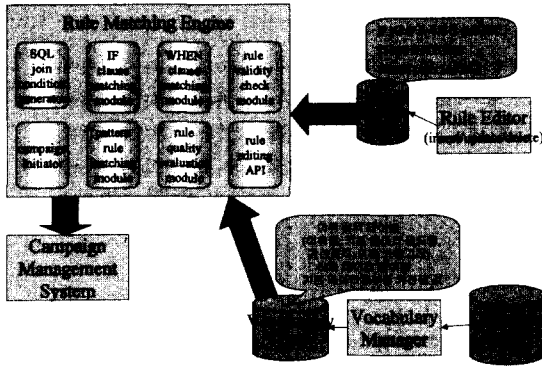


그림 1. CRM 환경에서의 룰 매칭 시스템

2.2 제안하는 CRM 룰의 구조

CRM 룰은 룰셋(rule set)단위로 구성된다. 룰셋이란 단일한 작업의 수행을 목적으로 하는 룰들의 집합이다. 룰 매칭 작업은 룰셋 단위로 이루어지기 때문에 같은 룰셋 내에 있는 룰들은 순서대로 수행된다. 이 때 룰셋 내의 모든 룰이 한꺼번에 수행되기 때문에 현재 시점에 수행하지 않아도 되는 룰도 함께 수행된다. 룰의 수행 시점 관련 조건을 분리하게 되면 현재 시점에 수행하지 않아야 하는 룰을 룰 수행 이전에 알아 낼 수 있다.

룰은 기본적으로 IF/THEN의 형태이다. 여기서 IF절은 주로 전체 고객 데이터로부터 이 룰이 대상으로 하는 고객집단을 선정하기 위한 조건이며, THEN절은 IF절의 조건을 만족하는 대상 고객집단에게 행할 룰 액션(rule action)으로서 CRM 환경에서의 룰 액션은 주로 이메일 발송 등과 같은 고객 대상 캠페인이다.

위와 같은 IF/THEN 형태의 룰을 이용하여 CRM 캠페인을 자동화하기 위해서는 룰 수행 시기를 명시적으로 표현할 수 있는 방법이 필요하다. 대부분의 CRM 룰 시스템에서는 룰 수행 시기에 관한 정보를 룰의 다른 조건들과 구분하지 않고 하나의 IF절로 표현한다. 각 조건의 카테고리, 즉 조건에서 사용하는 용어의 카테고리를 분류하지 않았기 때문에 시간요소조건을 룰의 다른 조건과 분리할 수 없는 것이다.

만일 시간요소조건이 IF절에 다른 카테고리의 조건들과 함께 들어가 있다면 다음과 같은 문제가 생긴다.

우선 카테고리가 다른 여러 조건들이 하나의 IF절 내에 섞여 있기

때문에 룰을 정의하거나 변경하는 사람이 룰의 의미를 구조적으로 파악하기가 어렵다. IF절은 고객 집단 선정에 관련된 조건을 의미하는 것으로 한정하여 하나의 IF절이 '고객'이라는 객체와 직접 혹은 간접적으로 관계가 있는 조건들로만 구성되도록 해야 의미가 명확해진다.

또한 위와 같이 IF절 내에 여러 카테고리의 조건들이 섞여 있으면 실제적으로 룰 매칭 시에 속도가 저하될 수 있다. 캠페인의 대상고객을 결정하는 부분인 IF절의 수행은 CRM 룰 매칭 프로세스에서 가장 중요한 부분이며 대용량의 고객 데이터 검색이 필수적이므로 가장 시간이 많이 걸리는 작업이라 할 수 있다. 룰 매칭 시간은 매칭해야 하는 룰의 갯수가 많아지고 고객 데이터의 규모가 커지게 되면 더욱 크게 증가할 것이다. 따라서 매칭할 룰의 개수를 줄일 수 있는 방법, 즉 고객 데이터 검색 횟수를 줄이는 방법이 필요하다. 이를 위해서는 모든 룰마다 고객 데이터를 검색하는 것이 아니라 룰 매칭을 필요로 하는 시점이 되었다고 판단되는 룰들에 한해서만 룰 매칭 작업을 수행함으로써 가능해진다. 이러한 판단을 내리기 위해서는 룰 매칭 시점에 관련되는 시간요소를 IF절로부터 분리하여 룰 내에 따로 저장해야 한다. 예를 들어 '크리스마스 시즌 or 토요일'이라는 시간요소조건이 IF절 내에 들어가 있다면 이를 IF절의 대상 고객집단 결정조건으로부터 분리하여 저장한 후 분리된 시간요소조건을 먼저 수행함으로써 크리스마스 시즌이거나 토요일이 아닌 날에는 이 룰의 IF절을 수행하지 않을 수 있다. 즉 이 룰로 인한 불필요한 고객 데이터 검색은 일어나지 않는다.

따라서 룰의 IF절과 THEN절 이외에 시간요소조건을 표현하기 위하여 WHEN절을 도입한다. 대상고객을 결정하는 부분을 IF절에 두고 시간요소조건을 WHEN절에 분리하여 저장하면 WHEN절이 만족되는 룰에 한해서만 IF절을 테스트하면 되므로 불필요한 전체 고객 데이터 검색을 막을 수 있고 이로 인하여 전체 룰 매칭 시간이 줄어들게 된다.

또한 대상고객 선정조건과 룰 수행 시점과 관련된 시간요소조건을 명확히 분리하였으므로 룰의 전체적인 의미를 구조적으로 파악할 수 있다. 따라서 룰의 생성과 유지, 보수 시에 사용자가 실수할 여지가 줄어들게 된다.

위와 같이 WHEN절을 IF절과 분리하기 위해서는 IF절에서 사용할 수 있는 어휘와 WHEN절에서 사용할 수 있는 어휘가 명확히 구분되어 있어야 한다. IF절과 WHEN절을 구분할 수 있는 방법은 조건에 사용한 어휘의 카테고리가 무엇인가에 달려 있기 때문이다. 즉 시간요소 카테고리의 어휘는 WHEN절에서만 쓸 수 있도록 하고, 고객과 직접 혹은 간접적으로 연관되는 어휘들은 IF절에서만 쓸 수 있도록 해야 한다. 만일 데이터베이스 필드를 IF절이나 WHEN절에서 바로 사용하게 되면 이러한 카테고리 구분이 불가능해진다. 따라서 룰에 사용하는 모든 용어는 어휘를 정의한 후에 그 어휘를 이용하도록 하여야 한다. 이에 관해서는 2.3에서 자세히 설명하겠다.

```

IF [대상고객집단 선정을 위한 조건]
THEN [캠페인 클래스(고객리스트+인자)]
WHEN [룰 실행 시점을 표현하는 시간 요소 조건]
    
```

그림 2. 시간요소조건을 분리 저장하기 위하여 WHEN절을 첨가한 쿼리의 구조

2.3 CRM 어휘

시간요소조건을 분리해내기 위해서는 어휘의 사용이 필요하다. 2.2에서 제시한 쿼리 구조에 따르면 IF절에 사용하는 용어는 고객과 직접 혹은 간접적으로 연관되는 것이어야 하고 WHEN절에 사용하는 용어는 쿼리 매칭 시점과 관련을 갖는 것이어야 한다. 따라서 어휘 정의 시에 그 어휘의 카테고리 정보를 입력하도록 하고 이 정보를 CRM 조건의 카테고리 구분의 기준으로 삼는다. 카테고리의 구분은 데이터베이스 필드를 조건에서 그대로 사용할 때에는 불가능하다. 따라서 어휘 정의의 테이블에는 기본 데이터 필드와 함께 어휘가 시간요소관련인지 고객관련인지는 나타내는 필드가 필요하다.

어휘의 사용은 쿼리 매칭 속도향상에도 관계된다. 쿼리의 조건에는 "고객.age > 30"과 같이 데이터베이스 필드를 직접 사용할 수 있다. 그러나 데이터베이스 필드 이름을 그대로 사용하는 경우가 아니라 데이터베이스의 데이터로부터 값을 유추하여 비교해야 하는 경우가 있다. 고객 테이블에는 생년월일의 형태로 기본 데이터가 들어가 있을 때 IF절에 나이를 비교하는 조건이 들어가는 경우가 그것이다. 이러한 조건을 테스트하기 위해서는 기본 데이터인 생년월일을 우리가 원하는 형태인 나이로 환산하는 계산 로직이 필요하게 된다.

계산 로직의 사용은 WHEN절에서 사용하는 시간요소조건에서는 필수적이다. WHEN절의 조건을 비교하기 위해서는 항상 현재 시점이 그 조건을 만족하는지를 테스트할 수 있어야 한다. "오늘이 크리스마스 시즌이면"이라는 조건이 있다고 하자. 이 때 몇 일부터 몇 일까지가 크리스마스 시즌인지를 정의해 놓은 데이터가 데이터베이스 내에 존재한다고 가정해더라도 이 데이터를 그대로 이용할 수는 없다. 이 조건을 테스트하기 위해서는 데이터베이스 내의 크리스마스 시즌 범위 정보와 오늘 날짜를 비교하여 그 범위에 속하는지를 판단하는 계산 로직이 개입되어야만 한다. 실제적인 조건 비교는 계산 로직의 결과로 생성된 데이터를 기반으로 이루어진다.

그러나 이러한 계산 로직을 쿼리 수행 시마다 호출하게 되면 쿼리 매칭 속도가 저하될 수 있다. 쿼리를 수행할 때마다 모든 고객 데이터를 읽어 들여 각 고객에 대하여 해당 계산 로직을 수행해야 하기 때문이다. 이것은 WHEN절의 조건을 검사할 때에도 마찬가지이다. 따라서 한 번 계산한 값을 해당 어휘 데이터 테이블에 저장하여 두고 어휘 데이터의 값이 변해야 할 경우에만 계산 로직을 다시 수행하게 한다면 쿼리 매칭 시간을 단축할 수 있다. 예를 들어 오늘이 크리스마스 시즌인지를 판단하는 계산 로직은 그와 관련되는 크리스마스 시즌 정의와 오늘 날짜, 이 두 가지 데이터가 변하는 순간에만 다시 수행하면 된다. 따라서 어휘 정의의 테이블에는 계산 로직 관련 정보가 들어가 있어야 하며 이 계산 로직 수행의 결과는 어휘 데이터 테이블에 저장되어 이후의 다른 쿼리 매칭 시에 사용함으로써 쿼리 매칭 속도가 향상될 수 있다.

어휘 정의 테이블	어휘 데이터 테이블
아래의 필드들을 포함한다.	위의 2)에 3)을 적용하여 나온 결과값을 저장한다.
1) 어휘명	
2) 기본 데이터 필드명	
3) 계산 로직 정보	
4) 어휘 카테고리 정보	

그림 3. 어휘 정의의 테이블과 어휘 데이터 테이블

3. 결론

본 논문에서는 CRM 쿼리에서 쿼리 매칭 시점과 관련한 시간요소조건을 분리하기 위하여 IF/THEN의 쿼리 구조에 WHEN절을 추가하였다. 시간요소조건을 분리를 통하여 쿼리의 의미를 구조적으로 파악하는 데에 도움을 줄 수 있으며 쿼리 매칭 시기 관련 시간요소조건을 먼저 테스트해 봄으로써 매칭 시기가 되지 않는 쿼리들을 선별해 낼 수 있고 실제 쿼리 매칭 작업은 WHEN절의 조건들을 만족하는 쿼리에 대해서만 행하면 되므로 전체 쿼리 매칭 시간이 줄어든다.

WHEN절과 IF절의 조건을 생성하기 전에 쿼리에 사용할 어휘를 정의하도록 함으로써 어휘 정의 시에 입력된 정보를 사용하여 WHEN절과 IF절을 구분할 수 있다. 또한 어휘 정의 시에 그 어휘 데이터를 생성하는 계산 로직 정보를 함께 저장하면 계산 로직의 결과를 어휘 데이터 테이블에 저장함으로써 쿼리 매칭 속도를 향상시킬 수 있다.

현재 본 논문에서 제시한 시간요소를 고려한 쿼리 모델을 적용하여 쿼리 매칭 엔진을 구축하고 있으며 향후에 쿼리 매칭 속도를 더욱더 향상시키기 위한 CRM 쿼리 모델 및 매칭 알고리즘의 개선 작업이 요구된다.

참고문헌

[1] Thomas Kahabka, Mari Korkea-aho, Gunther Specht, "GRAS : An adaptive personalization scheme for hypermedia databases". Proc. of HIM '97, pp. 279-292, 1997

[2] P.Resnick, et al., "GroupLens : An open architecture for collaborative filtering of netnews", Proc. of the ACM 1994 Conf. On CSCW, Chapel Hill, pp.175-186

[3] B.A.Huberman, et al., "Beehive : A system for cooperative filtering and sharing of information". Xerox Palo Alto Research Center, 1996, <http://www.parc.xerox.com/istl/groups/iea/abstracts/InformalOrganizations/beehive.html>

[4] Cassandra Millhouse, Colin Brash, Duncan Chapple, "E-CRM : Personalization Technologies for the Web", 2000, ISBN 1 902566-55-6

[5] Alex Berson, Stephen Smith, Kurt Thearling, "Building data mining applications for CRM", McGraw-Hill, pp. 4-14, ISBN 0-07-134444-6