

# 스마트폰 음성 인식 긴급전화 서비스



황차동 아이티엘 대표

## 1. 머리말

최근 심심치 않게 발생하고 있는 붕괴, 충돌, 화재 및 납치 등의 사고 소식으로 국민들은 사건·사고가 단순하게 뉴스 혹은 신문의 헤드라인을 넘어서서 실제 '나'의 삶에 발생할 수 있는 일로 체감하고 있다. 이는 자연스럽게 안전과 더불어 실제 사고가 발생하여 긴급한 상황에 처했을 때의 대응방법에 대한 관심으로 이어지고 있다.

2014년부터 TTA에서는 국민 편익과 실생활에 도움을 줄 수 있는 『국민 편익/생활 밀착형 표준』을 추진하고 있으며, '스마트폰 음성 인식 긴급전화 서비스' 표준은 이러한 '안전'과 '대응방법'에 대한 국가적, 국민적 니즈(Needs)에 발맞춘 생활표준 제정의 첫 디딤돌이다.

본 표준인 '스마트폰 음성 인식 긴급전화 서비스'는 사건·사고로 인하여 발생하는 인적피해를 줄이기 위하여 태동하였다. 특히, 사고 발생의 긴급한

상황에서 신속하고 효과적으로 구조요청을 할 수 있는 방법을 고안하던 중에 '음성인식'과 '스마트폰'에서 그 해결책을 찾았다.

현재 한국의 스마트폰 보급률은 67.6%로 전 세계 국가 중 가장 높은 것으로 나타났다. 즉, 10명 가운데 6명 이상이 스마트폰을 소유하고 있으며 이로써 스마트폰이 이미 우리 생활에 밀접하게 연결되어 있는 필수품 중 하나라는 것은 의심할 여지가 없다. 만약 어느 날 갑자기 사고가 발생하여 긴급한 상황에 처하게 되었을 때, 스마트폰을 이용하여 구조요청을 할 수 있다면 국민 10명 가운데 6명 이상이 이러한 서비스를 통하여 위급한 상황을 모면할 수 있을 것이다.

물론 이미 스마트폰을 이용한 긴급전화 서비스는 존재한다. 하지만 본 표준의 가장 큰 핵심은 '음성인식'이다. 기존의 스마트폰을 통한 긴급전화 서비스의 경우, 애플리케이션을 실행해야 하거나 해당 버튼을 눌러서 관련 프로그램을 실행해야 하는 것

과 같은 부수적인 절차들이 필요하다. 하지만 본 표준은 ‘음성 인식’을 이용하기 때문에 신체적으로 제한된 상황에서 부수적인 절차 없이 음성만으로 긴급전화 서비스를 가능하게 한다. 즉 건물 붕괴 등으로 손을 사용할 수 없는 상황에 직면하거나 어떤 특수한 상황에 의해 결박되어 있는 상태에서 긴급 전화를 필요로 할 때는 ‘음성 인식’만으로 통화가 되어야 하는 것이다.

본 표준의 제정을 위하여 2014년 1월 ‘음성인식 긴급전화 표준화 전담반’이 구성되어 표준의 시나리오 및 기술적 요구 사항 등에 관한 협의가 진행되었다. 초기에는 통신 3사, 제조 3사 그리고 통신 관련 주요 기관 및 기업들로 구성되어 음성인식 긴급전화 서비스의 구현을 위한 서비스 시나리오를 구축하였다. 가장 기본적이면서 뼈대가 되는 본 서비스의 시나리오를 합의하는 데 있어서 통신 3사와 제조 3사의 의견 합의는 쉽지만은 않았다. 특히, 본 표준의 경우에는 긴급한 상황에서 서비스 사용자의 ‘안전’과 ‘구조’라는 목적을 가지고 있기 때문에 긴급구조기관 및 경찰청의 의견과 실제 사용하게 될 국민(소비자)의 의견 역시 굉장히 중요한 부분을 차지하였고 이로 인하여 의견 합의의 어려움은 더 컸다. 뿐만 아니라, 신속하고 효율적인 서비스의 구현만큼이나, 서비스의 오작동 부분을 고려해야 하기 때문에 시나리오의 세부절차와 함께 구체적인 기술적 요구사항의 표준화를 협의하는 데 있어서 물리적인 시간이 많이 투입되었다. 그러나 ‘음성인식 긴급전화 표준화 전담반’의 각계 각층 위원들이 본 서비스에 대한 높은 관심과 함께 적극적으로 지원하여 2014년 9월 본 제안의 목적과 부합되는 최적의 서비스 표준화를 이끌어냈다. 이에 따라, 국민의 편익을 위하여 제안된 ‘스마트폰 음성 인식 긴급전화 서비스’ 표준에 관하여 소개하고자 한다.

## 2. 주요내용

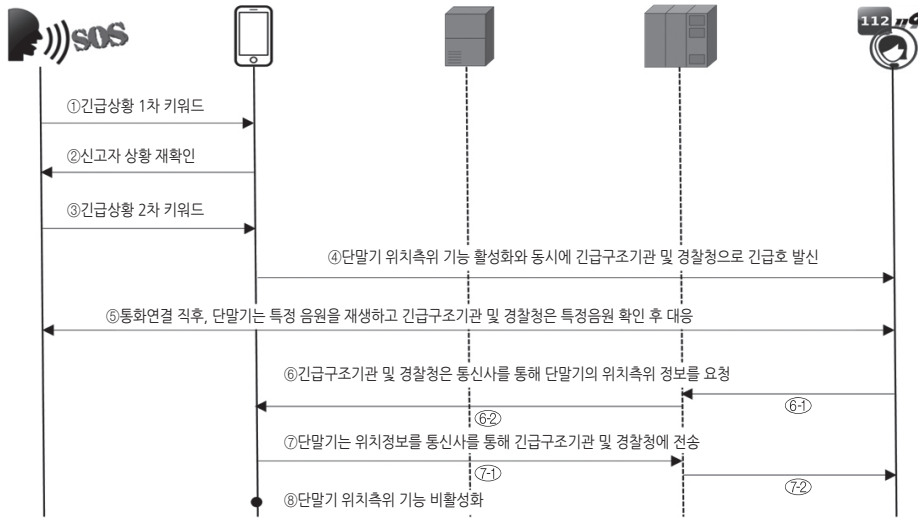
### 2.1 ‘스마트폰 음성 인식 긴급전화 서비스’ 구현을 위한 표준의구성및 범위

본 표준은 스마트폰을 이용하여 사용자의 구조 요청부터 긴급구조기관 및 경찰청과 통화 연결까지의 시나리오 구현을 위해 필요한 기술적 요구사항을 명시한 것으로 전체 시나리오와 음성인식, 긴급전화 음성 키워드 및 사용자 정보 전송 부분에 대한 세부 기술적 규격을 ‘재난·사고 상황’과 ‘범죄 상황’으로 구분하여 제시한다.

### 2.2 ‘스마트폰 음성 인식 긴급전화 서비스’ 구현을 위한 시나리오

재난, 사고 혹은 범죄와 같이 긴급한 상황에 처한 사용자의 구조 요청 시점부터 실제 긴급구조기관 및 경찰청과의 통화 연결까지의 서비스 제공을 위한 시나리오이며, 사용자와 스마트폰 간에 긴급한 상황이 맞는지에 대한 판단 확인 절차가 포함되어 있다. [그림 1]은 ‘음성 인식 긴급전화 서비스 시나리오’를 보여준다. 서비스 흐름도의 절차별 세부내용은 다음과 같다.

- ① 양손을 사용할 수 없는 긴급상황에 처한 사용자가 ‘1차 음성키워드’로 스마트폰 단말을 대기 모드에서 동작모드로 전환
- ② 스마트폰은 긴급 상황이 맞는지의 판단을 위해 ‘2차 음성 키워드’ 입력 대기
- ③ 사용자는 ‘2차 음성 키워드’로 긴급 상황임을 스마트폰에 인식
- ④ 스마트폰은 위치 측위 기능(GPS, WiFi)을 활성화시키고, 긴급구조기관 및 경찰청으로 긴급호 발신



[그림 1] 긴급 상황 음성 인식 긴급전화 서비스 흐름도

- ⑤ 긴급 호 연결 직후 → 스마트폰 단말이 특정 음원<sup>1)</sup>을 긴급구조기관 및 경찰청에 재생 → 사용자와 응답원 간 통화 시작
- ⑥-1 긴급구조기관 및 경찰청은 통신사 서버로 스마트폰 단말의 위치 측위 정보 요청
- ⑥-2 통신사 서버는 통신사 교환기를 거쳐 스마트폰 단말의 위치 측위 정보 획득
- ⑦-1 스마트폰 단말은 통신사 교환기를 거쳐 통신사 서버로 위치 측위 정보 제공
- ⑦-2 통신사 서버는 이 정보를 긴급구조기관 및 경찰청에 제공
- ⑧ 통화 종료 후, 스마트폰 단말의 위치 측위 기능(GPS, WiFi)을 원 상태로 복귀시키거나 비활성화시킴

는 일련의 절차들 안에서 긴급상황 판단 절차를 2단계 포함하고 있어 오작동 및 오신고 확률을 줄일 수 있다. 또한, 통화연결 시에 ‘특정 음원’을 가장 먼저 재생하는 절차를 넣어서 긴급구조요청을 하는 사용자가 ‘긴급한 상황’과 ‘신체적 제한 상황’이라는 2가지 정보를 제공하여 응답원의 대응을 고려하였다.

### 2.3 ‘스마트폰 음성 인식 긴급전화 서비스’ 구현을 위한 기술적요구사항

서비스 구현을 위하여 음성인식, 긴급전화 음성 키워드 및 사용자 정보 전송 부분에 대한 세부 기술적 규격을 ‘재난·사고 상황’과 ‘범죄 상황’으로 구분하여 명시하고 있으며, 이는 <표 1>, <표 2>, <표 3>으로 제시하였다.

본 표준은 긴급 상황에 대해 ‘음성 인식’으로 대응하기에 주변의 소음 대비 사용자의 음성을 인식

특히, ‘긴급 상황 음성 인식 긴급전화 서비스 흐름도’

1) 특정 음원: ‘음성 인식 긴급전화 서비스’를 통하여 연결된 전화라는 알림 음성 멘트로서, 긴급구조기관 및 경찰청 응답원이 전화를 처음 수신하는 시점에서 재생되는 음원

<표 1> 긴급상황별 기술적요구 사항

구분	대상	요구사항	
		재난·사고 상황	범죄 상황
음성	스마트폰이 응답해야 하는 음성의 크기(마이크 도달 기준)	70dB SPL	
	키워드의 길이	4음절 이상, 600ms 이상(권고: 5음절 이상, 1초)	
	주변 소음대비 음성 SNR	12dB	
	음성 인식 방법	화자 독립	화자 독립(default) or 화자 종속
사용자 정보 전송	사용자 정보의 전송 내용	위치정보(기지국, GPS, WiFi 정보)	
특정음원	음원	예) '음성긴급전화'	
신고자 상황확인	스마트폰이 긴급 상황 여부를 확인하기 위한 알림	화면밝기, 신호음, 진동 등	

<표 2> 음성 키워드

구분	재난·사고 상황	범죄 상황	
음성 키워드 저장 방식	스마트폰 내부 메모리		
음성키워드 개수	복수		
2차 음성 키워드 선택 방법	프리-어사인드(pre-assigned)	프리-어사인드 (pre-assigned)	셀프-리코디드 (self-recorded)
음성 인식 방법	화자 독립	화자 독립	화자 종속
1차 음성키워드	소방서(default)	경찰서(default)	
2차 음성키워드	도와주세요(default), 살려주세요, 사람살려요, 소방서 연결, 긴급전화, 비상전화, 긴급통화, 신고전화	도와주세요(default), 살려주세요, 사람살려요, 경찰서 연결, 긴급전화, 비상전화, 긴급통화, 신고전화	

<표 3> 범죄 상황에서 음성 키워드 선택 방법

2차 음성 키워드 선택 방법	음성 인식 방법	비고
프리-어사인드(pre-assigned)	화자 독립	default
셀프-리코디드(self-recorded)	화자 종속	-

하는 기술을 세분화하여 요구사항을 명시하였다. 그리고 '음성 인식'만큼이나 실질적으로 구조에 있어 가장 중요하게 이용되는 사용자 정보에 관해서는 '긴급구조기관 및 경찰청'의 의견과 '제조 3사'의 의견을 반영하고 특히, 관련 법규를 고려하여 우선 현재는 위치정보(기지국, GPS, WiFi)로 한정하여 제시하였다.

기술적 요구사항을 보면 2개의 긴급상황을 구분하여 제시하고 있다. 범죄 상황의 경우, 재난·사고 상황과는 달리 제3자가 존재할 수 있기 때문에 음성

인식 부분에서도 목소리 구별이 가능한 '화자 종속' 방법을 추가로 구성하여 좀 더 복잡한 범죄 상황으로부터 사용자의 '안전'을 확보하기 위함이다.

음성 인식 긴급전화 서비스의 '음성 키워드'는 사용자가 긴급상황에 처하고 처음 구조요청을 하는 '1차 음성 키워드'와 긴급상황이 맞는지의 여부를 스마트폰과 사용자 사이에서 확인하는 '2차 음성 키워드'로 구성된다.

음성키워드는 스마트폰 제조단계에서 내부 메모리에 저장(pre-loaded)하는 방식을 채택하였으며,

개수는 복수개로 합의되었다.

1, 2차 음성키워드의 경우, 각 상황에 맞게 1개의 기본 키워드(default)를 설정해 두었지만, 사용자의 기억이 용이한 키워드가 다름을 고려하여 몇 개의 키워드 군을 제시하였고 추가제공이 가능하다. 하지만 음성인식에 있어 기술적인 제한이 있기에 최소 4음절 이상, 600ms 이상의 발성시간, 5음절 이상, 발성시간 1초 이상의 권고사항을 명시하였다. 또한 필요에 따라 음성 키워드는 갱신될 수 있고 상세한 갱신방법과 주기는 단말기 제조업체에 자율성을 부여하였다. 다만, 기 설정된 음성키워드의 작동에 문제가 없어야 함을 명시하였다.


특히, 제3자의 존재 여부에 따른 사용자의 '안전' 확보는 2차로 스마트폰과 사용자 간의 긴급상황 확인절차 시에 가장 필요하다. 이러한 필요를 해결하고자 본 표준에서는 '2차 음성 키워드'의 선택에 있어서 2가지 방법을 제시하고 있다. 첫 번째는 재난·사고 상황처럼 제시된 다수의 키워드 군에서 사용자가 선호하는 키워드를 설정하는 프리-어사인드(pre-assigned) 형태로 화자독립방식으로 사용자가 음성인식을 해야 한다. 두 번째는 사용자가 다수의 키워드 군에서 선호하는 키워드를 직접 자신의 음성으로 녹음하여 설정하는 셀프-리코디드(self-recorded) 형태로 '화자 종속' 방식으로 사용자의 음성이 인식된다. 이는 제3자가 함께 있는 상황에서도 스마트폰이 녹음된 목소리만을 구별하여 인식함으로써 위험한 상황 속에서 안전을 확보할 수 있다.

### 3. 맺음말

10월 제정된 '스마트폰 음성 인식 긴급전화 서비스' 표준은 통신 3사, 제조 3사 및 통신 관련 주요 기관 및 기업뿐만 아니라 실제 이용대상자인 긴급구조기

관 및 경찰청과 사용자에게 해당하는 소비자 단체의 의견을 모두 수렴하여 완성되었다.

또한, 표준제정에 그치지 않고 2015년 지자체 중심 단계별 시범서비스 검증을 계획하고 있으며, 그 이후에는 실제로 단말기에 장착하여 서비스를 지원할 계획을 가지고 있다. 그만큼 본 표준은 활자로만 끝나지 않고 실생활에서의 구현을 목표로 하고 있다.

이처럼 각계각층에서 관심을 가지고 '스마트폰 음성 인식 긴급전화 서비스' 표준의 실제 구현에 관심을 갖고 지원하는 데는 그 중요성을 체감하고 있기 때문이라 판단된다. 우선 첫 번째는 손을 사용할 수 없어도 전화통화를 연결하여 구조요청을 할 수 있는 것과 같이 최근에 빈번히 발생하고 있는 사건·사고의 긴급한 상황에서 새로운 채널을 추가하여 조금이나마 인적 피해를 줄이는 데 있다. 두 번째는 연일 헤드라인을 장식했던 '안전 불감증, 대한민국'에 대한 하나의 방안으로 '안전'과 '대응'에 대한 인식 확대에 있다. 그리고 세 번째로는 긴급한 상황이 벌어져 실제 구조가 진행될 때, 스마트폰 위치정보를 이용하여 신고자의 위치를 보다 정확히 파악하여 구조시간을 단축시키는 데 있다. 하지만 무엇보다도 '국민 편익/생활 밀착형 표준'의 첫 디딤돌이라는 점에 있다. 실생활에 도움이 필요한 부분에 대하여 각계각층이 함께 아이디어를 모으고 표준을 제정함으로써 '표준화'인식 확산에 크게 기여할 것이며, 이를 상용화하여 국민들이 실제로 제정된 서비스를 이용하게 된다면 국민의 편익 확대에 이바지할 것이라 기대한다. 

### [참고문헌]

[1] TTA, '스마트폰 음성 인식 긴급전화 서비스(TTAK.KO-06.0367)', 2014. 10. 13.