

TOOLS THU THẬP DỮ LIỆU KHÁCH HÀNG TỪ MẠNG XÃ HỘI, ỨNG DỤNG CHO CÁC DOANH NGHIỆP VỪA VÀ NHỎ

Phạm Thị Liên*, Trần Tuấn Việt,
Nguyễn Quang Hiệp, Nguyễn Thu Phương, Trần Thị Tuyết
Trường Đại học Công nghệ thông tin & Truyền thông – ĐH Thái Nguyên

TÓM TẮT

Sự cộng hưởng trong những năm vừa qua của các công nghệ số có nhiều đột phá (như điện toán đám mây, internet vạn vật, dữ liệu lớn, trí tuệ nhân tạo...) đã báo hiệu những thay đổi lớn lao đang bắt đầu xảy ra, được gọi ở nhiều nơi là cuộc cách mạng công nghiệp lần thứ tư. Cùng với cuộc cách mạng 4.0, với sự phát triển mạnh mẽ của thương mại điện tử đã đẩy mạnh các hoạt động kinh doanh, buôn bán, quảng bá sản phẩm, dịch vụ diễn ra trên khắp các kênh thông tin xã hội, đặc biệt là mạng xã hội facebook. Người dùng chia sẻ ý kiến nhận xét, đánh giá, phản hồi của mình về sản phẩm, ngành hàng trên mạng xã hội. Còn các doanh nghiệp thông qua đó sẽ có cơ hội thấu hiểu khách hàng của mình, biết được chủ đề mà họ quan tâm trên mạng xã hội là gì; từ đó đưa ra được các chiến lược kinh doanh phù hợp, đạt hiệu quả. Do đó, chúng tôi xây dựng công cụ hỗ trợ thu thập dữ liệu với khả năng thu thập dữ liệu thông minh, kịp thời, phân loại dữ liệu cần thiết, là giải pháp tối ưu cho doanh nghiệp vừa và nhỏ ở Việt Nam trong thời đại công nghệ số.

Từ khóa: facebook; mạng xã hội; online ads; khai phá dữ liệu mạng xã hội; chatbot.

Ngày nhận bài: 30/11/2019; Ngày hoàn thiện: 25/5/2020; Ngày đăng: 29/5/2020

TOOLS COLLECT CUSTOMER DATA FROM SOCIAL NETWORKS, APPLICATIONS FOR SMALL AND MEDIUM ENTERPRISES

Pham Thi Lien*, Tran Tuan Viet,
Nguyen Quang Hiep, Nguyen Thu Phuong, Tran Thi Tuyet
TNU - University of Information and Communication Technology

ABSTRACT

Resonance in recent years of digital technologies has breakthroughs (such as cloud computing, internet of things, big data, artificial intelligence ...) has signaled that great changes are beginning to happen. out, known in many places as the fourth industrial revolution. Along with the 4.0 revolution, with the strong development of e-commerce makes the promotion of business, trade and promotion of products and services taking place throughout the Social information channels, especially facebook social network. Users share their opinions, comments, reviews about products and industries on social networks. And businesses through that will have the opportunity to understand their customers, know what topics they are interested in on social networks. Since then, we have come up with appropriate and effective business strategies. Therefore, we build a data collection support tool with the ability to collect data intelligently, promptly, and classify necessary data, which is the optimal solution for small and medium enterprises in Vietnam in the digital technology.

Keywords: advertising; facebook; Social Network; marketing; online ads; data mining social network; chatbot.

Received: 30/11/2019; Revised: 25/5/2020; Published: 29/5/2020

* Corresponding author. Email: ptlien@ictu.edu.vn

1. Giới thiệu

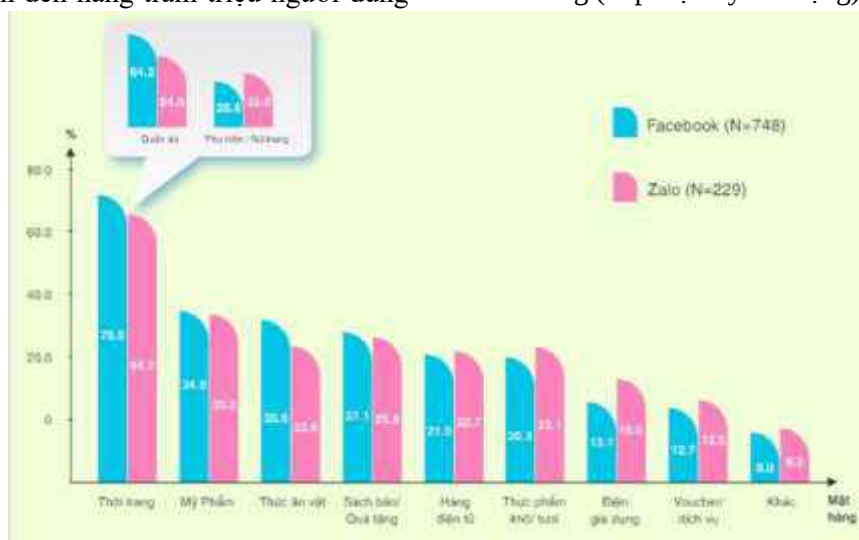
Mạng xã hội là nơi kết nối mọi người với nhau để nói chuyện, chia sẻ tâm tư, sở thích và kết nối với nhau thông qua phương tiện truyền thông xã hội hay còn gọi là “Social Media”. Các trang web truyền thông mạng xã hội chứa nội dung do hàng trăm hoặc thậm chí hàng triệu người khác nhau tạo ra. Mỗi người dùng đều thông qua mạng xã hội để thể hiện ý kiến nhận xét của mình, cái nhìn của mình về một sự kiện nào đó; không những thế, họ còn thể hiện rõ sự yêu/ghét của mình đối với những nội dung được người khác chia sẻ trên mạng xã hội. Thông qua mạng xã hội, các cộng đồng mạng có liên quan có thể được liên kết để giúp người dùng tìm thấy các địa chỉ liên hệ có thể hữu ích với họ. Mạng xã hội không chỉ là nơi để mọi người giao lưu, kết nối và chia sẻ, mạng xã hội còn là “mảnh đất vàng” cho kinh doanh online [1], hỗ trợ tìm kiếm khách hàng tiềm năng, tương tác, thậm chí là quảng cáo, giúp xây dựng thương hiệu doanh nghiệp hoặc xây dựng thương hiệu cá nhân. Có rất nhiều mặt hàng được giới thiệu đến người tiêu dùng thông qua mạng xã hội (hình 1).

Hiện nay, các trang mạng xã hội phổ biến với số lượng lên đến hàng trăm triệu người dùng

là: Instagram, Facebook, Tumblr, Twitter, Pinterest, Snapchat, Youtube, LinkedIn,...

Theo thống kê của Wearesocial.com [2], thì đến tháng 1 năm 2018, Việt Nam có 96,02 triệu người dân, trong đó có 55 triệu người dùng Social Media (chiếm 55%), 35 triệu người sở hữu điện thoại thông minh (chiếm 36%). Trên các website, diễn đàn, mạng xã hội ở Việt Nam cũng thu được hơn 50 triệu bài posts, comments, shares mỗi ngày (hệ thống phân tích SocialHeat, YouNet Media, 2018). Tuy nhiên, khi mà ngày càng nhiều người tiêu dùng sử dụng mạng xã hội kèm theo hành vi mua sắm online. Chỉ cần sử dụng thiết bị thông minh với vài thao tác lướt và chạm, thì giờ đây người tiêu dùng có thể truy cập vào nhiều nguồn tin, nhiều trang bán hàng. Đồng thời, họ cũng có nhiều cách để phản hồi về chất lượng của sản phẩm cũng như các dịch vụ mà doanh nghiệp cung cấp.

Các báo cáo thống kê, phân tích về hành vi, sở thích của khách hàng là cơ sở để duy trì niềm tin của khách hàng, thực hiện khuyến khích khách hàng mua sản phẩm, mời khách hàng mua các sản phẩm liên quan, hoặc thúc đẩy tiếp thị theo hình thức word-of-mouth marketing (tiếp thị truyền miệng).



Hình 1. Những mặt hàng được mua sắm qua facebook/zalo

Để có đủ dữ liệu phân tích cho các hoạt động của người tiêu dùng, thì dữ liệu được thu thập thông qua giao tiếp với khách hàng, qua hệ thống quản lý quan hệ khách hàng hoặc thống kê trên website là không đủ. Các thảo luận của khách hàng cần được lắng nghe ở khắp mọi nơi trên

internet. Việc sở hữu cho riêng mình một hệ thống thu thập, phân tích dữ liệu đa kênh sẽ là giải pháp chủ đạo cho các doanh nghiệp và thương hiệu trong thời gian tới. Tận dụng dữ liệu đó để tối ưu hóa quảng cáo online và quảng bá thương hiệu. Quá trình này bao gồm việc thu thập, phân tích dữ liệu từ khách hàng như thói quen, hành vi, mong muốn, hoàn cảnh... để tối ưu hóa hiệu quả các chiến dịch marketing. Với sự phát triển của công nghệ, việc có được dữ liệu khách hàng hiện nay đã không còn nhiều khó khăn. Doanh nghiệp có thể thu thập được qua giao tiếp với khách hàng, qua các chiến dịch quảng cáo trực tuyến, từ dữ liệu nhân khẩu học; dữ liệu từ hệ thống quản lý quan hệ khách hàng (Customer Relationship Management -CRM) hoặc có thể mua dữ liệu từ bên thứ 3.

Xác định đúng mục đích và tối ưu hóa nguồn dữ liệu khổng lồ (big data) sẽ mang đến rất nhiều lợi ích về nhiều mặt khi thực hiện chiến dịch marketing như: hiệu quả bán hàng, trải nghiệm người dùng đối với dịch vụ, hiệu quả ROI marketing, lợi nhuận, tăng trưởng kinh doanh, độ nhận diện thương hiệu, nguồn vốn đầu tư... Do đó, các doanh nghiệp nên thay đổi tư duy marketing [3], thay vì đầu tư để tìm kiếm khách hàng mới thì nên phát triển, thu thập dữ liệu để chăm sóc những khách hàng đã có. Nhờ vậy, doanh nghiệp sẽ có được nhiều khách hàng tiềm năng, xây dựng hình ảnh thương hiệu tốt và thu hút thêm khách hàng mới thông qua uy tín đã tạo dựng được.

2. Phương pháp thu thập dữ liệu trên mạng xã hội

Với tầm quan trọng của dữ liệu khách hàng trong thời đại công nghệ số. Nghiên cứu thu thập dữ liệu khách hàng trên mạng xã hội đã và đang trở nên rất thiết thực. Đã có rất nhiều các nghiên cứu về phương pháp thu thập dữ liệu trên các mạng xã hội cũng như các hoạt động thu thập dữ liệu của các công cụ lắng nghe và theo dõi mạng xã hội. Social Listening [4] là một mô hình kinh doanh khác của ngành nghiên cứu thị trường. Các công cụ thu thập dữ liệu sẽ cho biết phản hồi của người tiêu dùng về thương hiệu, cho biết người tiêu dùng quan tâm đến gì trong ngành

hàng của mình từ đó đưa ra được chiến lược quảng bá nhanh và chi phí thấp.

Giải pháp Social Listening có thể theo dõi tất cả các bài viết được đăng tải trên mạng xã hội có ảnh hưởng lớn và tất cả các fanpage của Facebook, Google+ và Youtube và có thể đưa ra cảnh báo kịp thời trước khi một vấn đề tiêu cực về thương hiệu bị phát tán và trở thành khủng hoảng.

Hiện nay có rất nhiều nhà cung cấp giải pháp Social Listening. Ở các nước phát triển có rất nhiều công cụ Social Media Listening, nhưng phần lớn chỉ hỗ trợ tiếng Anh và các ngôn ngữ của các nước phát triển. Một số hệ thống tiêu biểu là Radian 6, Sysomos, Meltwater, NetBase... nhưng phí sử dụng rất cao. Một số hệ thống có bản dùng thử (cũng cho tiếng Anh) như Social Mention, Topsy, Sprout Social...

Tuy nhiên, các hệ thống *social listening* do châu Âu, Mỹ phát triển thường không thể thu thập được hết các nguồn thông tin của các quốc gia dùng ngôn ngữ riêng như Việt Nam do sự khác biệt về ngôn ngữ và sự tập trung nguồn lực đầu tư. Do đó các hệ thống này thường không thể áp dụng tốt tại các nước có sử dụng ngôn ngữ riêng như Việt Nam, Thái Lan, Malaysia...

Ở Việt Nam, ngoài Noti5 là hệ thống thu thập tin trực tuyến của ePi (sau đó ePi hầu như đã từ bỏ) thì còn có hệ thống SocialHeat (của YouNetMedia), Boomerang, Buzzmetrics.

Buzzmetrics [4] là một công cụ đang được dùng để theo dõi các chiến dịch và thương hiệu quản lý bởi các tập đoàn lớn Samsung, Coca-Cola, Unilever và các Agencies global như Ogilvy, Maxus, Leo Burnett, Phibious, Performics... tại Việt Nam và được đánh giá cao.

Ở Việt Nam, một số trang mạng xã hội phổ biến để doanh nghiệp vừa và nhỏ có thể nghiên cứu là Facebook, diễn đàn, trang Tin tức, Review site và website của các nhà bán lẻ trực tuyến. Nhược điểm của các hệ thống thu thập dữ liệu là tốc độ cập nhật dữ liệu. Một số hệ thống Social Listening tại Việt Nam hầu hết không đáp ứng được yêu cầu về tốc độ cập nhật dữ liệu theo thời gian thực.

Vậy giải pháp nào cho các doanh nghiệp vừa và nhỏ ở Việt Nam? Từ những nghiên cứu và phân tích đã đưa ra, nghiên cứu này của chúng tôi nhằm xây dựng một công cụ thu thập dữ liệu trên mạng xã hội facebook, hỗ trợ cho các doanh nghiệp vừa và nhỏ ở Việt Nam. Với dữ liệu thu thập được doanh nghiệp sẽ biết được những đánh giá/ phản hồi của người dùng về sản phẩm, thương hiệu của mình, cũng như biết được những thông tin nổi bật/ xu hướng về sản phẩm mà người dùng đang quan tâm; từ đó hoạch định được chiến lược phát triển, kinh doanh của doanh nghiệp mình.

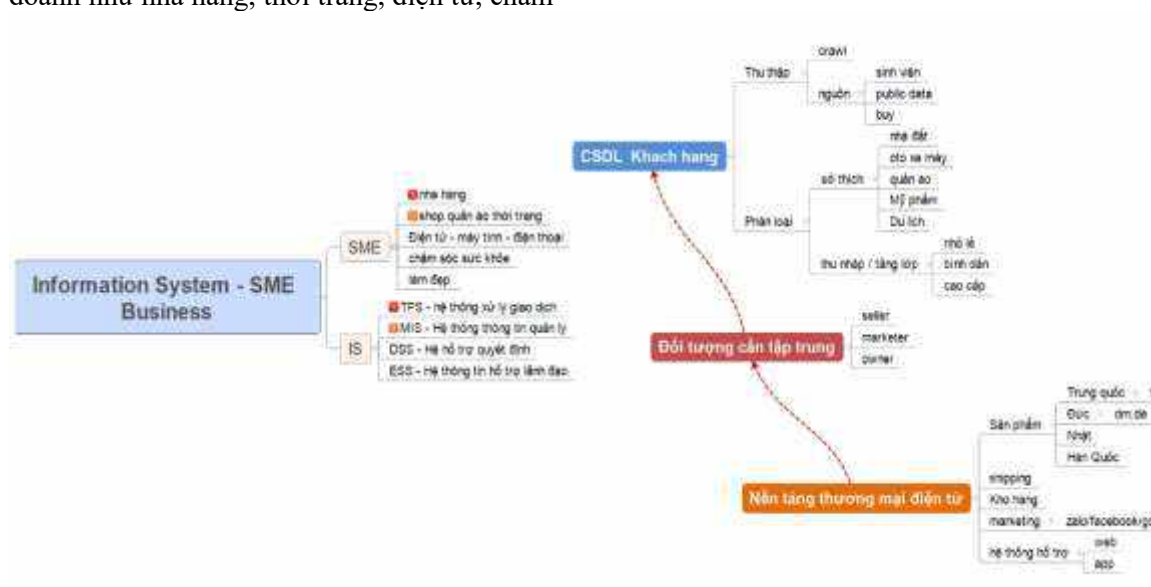
3. Triển khai xây dựng hệ thống thu thập dữ liệu

3.1. Khảo sát và phân tích mô hình kinh doanh vừa và nhỏ tại Việt Nam

Chúng tôi đã khảo sát mô hình kinh doanh của một số doanh nghiệp vừa và nhỏ (Small Medium Enterprise - SME) tại Việt Nam và đưa ra cái nhìn tổng quan về mô hình hệ thống kinh doanh của doanh nghiệp. Mô hình SME ở Việt Nam gồm một số lĩnh vực kinh doanh như nhà hàng, thời trang, điện tử, chăm

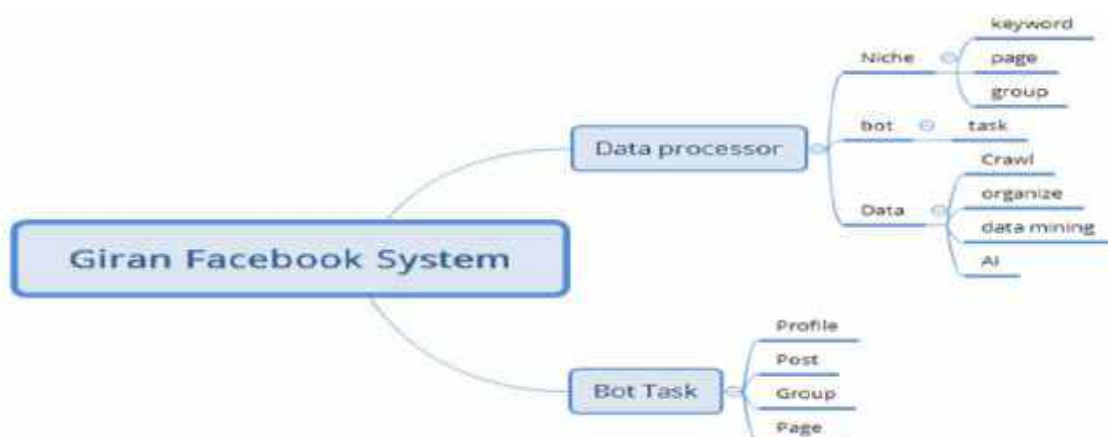
sức khỏe, làm đẹp. Hệ thống thông tin (IS) đối với mô hình SME gồm hệ thống xử lý giao dịch, hệ thống thông tin quản lý, hệ hỗ trợ quyết định, hệ thông tin hỗ trợ lãnh đạo. Để thúc đẩy sự phát triển của SME, dựa trên nền tảng của thương mại điện tử, chúng tôi đưa ra kế hoạch thực hiện: Đối tượng mà chúng tôi quan tâm là người bán, người tiếp thị, chủ sở hữu. Nhiệm vụ trọng tâm là thu thập và phân loại cơ sở dữ liệu của khách hàng (hình 2).

Sau khi khảo sát thực tế tại doanh nghiệp chúng tôi đã phân tích nhu cầu cũng như tình hình thực tế của một số doanh nghiệp, từ đó biết được mục đích của các doanh nghiệp, hình thức kinh doanh, nhu cầu sử dụng các dữ liệu trên facebook. Thống kê các nguồn lực của doanh nghiệp về đội ngũ nhân viên, kỹ năng sử dụng công nghệ, hệ thống cơ sở hạ tầng, phần cứng của doanh nghiệp. Từ đó đưa ra mô hình của hệ thống thu thập dữ liệu trên facebook phù hợp với SME tại Việt Nam. Đáp ứng được mục đích của doanh nghiệp cũng như chi phí đầu tư là thấp nhất (hình 3).

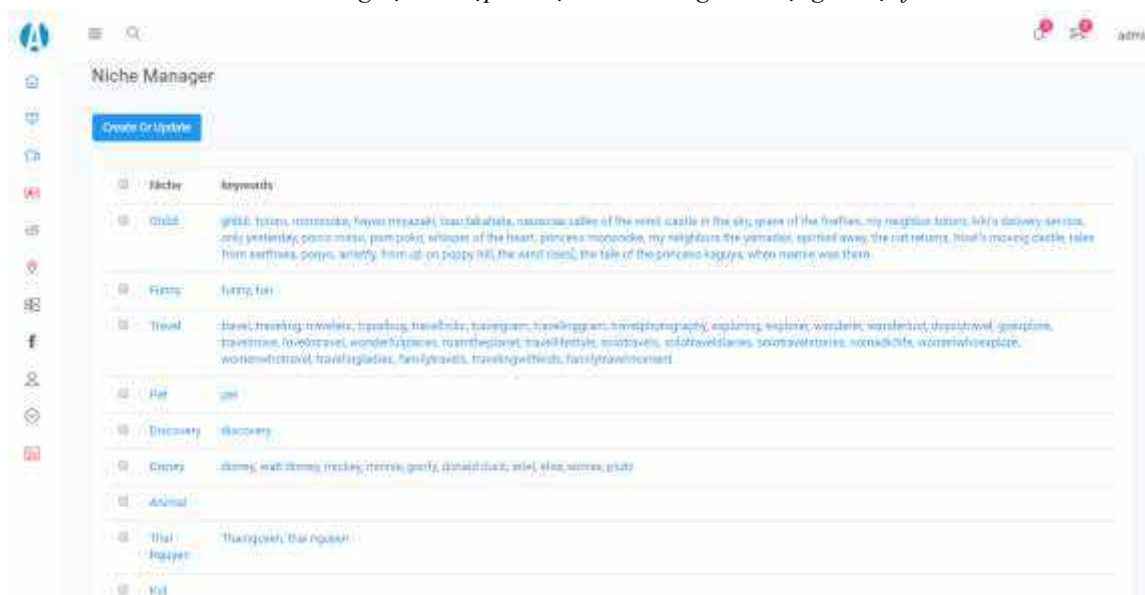


Hình 2. Mô hình kinh doanh của các doanh nghiệp vừa và nhỏ ở Việt Nam

Hệ thống thu thập dữ liệu mà chúng tôi xây dựng cơ bản gồm 3 phần. Phần thứ nhất, chúng tôi xây dựng các **Niche** để quản lý thu thập dữ liệu từ các fanpage và groups trên facebook. Thứ hai, chúng tôi xây dựng các **Bot, Crawl** nhằm mục đích tham gia vào các trang để theo dõi và lấy thông tin người dùng từ các trang cũng như thu thập các dữ liệu cần thiết khác. Đây là phần quan trọng nhất của hệ thống. Thứ 3 là xử lý dữ liệu thu thập được.



Hình 3. Sơ đồ công cụ thu thập dữ liệu khách hàng trên mạng xã hội facebook



Hình 4. Thu thập dữ liệu từ các sites được liệt kê

3.2. Các phương pháp chính để thu thập dữ liệu

Có 2 phương pháp chính: Từ khóa (Keywords) và Trang (Sites) [6], [7].

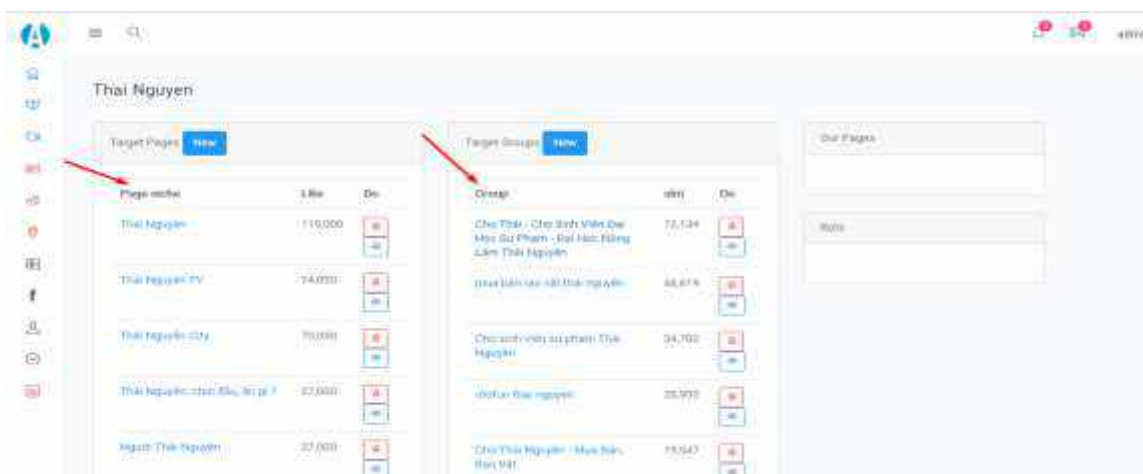
Phương pháp 1: Thu thập dữ liệu bằng keywords

Phương pháp này được áp dụng đối với các mạng xã hội toàn cầu như Facebook, Google Plus, Youtube, Twitter, Instagram... Phương pháp thực hiện chính là công cụ sẽ được kết nối với các API (Application Programming Interface) của các mạng xã hội và gửi yêu cầu cho phép lấy dữ liệu có chứa keywords được chỉ định. Phương pháp này theo nguyên tắc cho phép lấy dữ liệu của toàn bộ mạng xã hội, bao gồm cả các thông tin cá nhân, nhưng

trên thực tế nó lại phụ thuộc vào sự hạn chế của các mạng xã hội này. Ví dụ, với Facebook có sự hạn chế về số lượt người có khả năng tiếp cận tự nhiên với các bài viết của các chủ fanpage và các nhà quảng cáo, Facebook cũng không trả lại đầy đủ và nhất quán các bài viết cá nhân cho các công cụ thu thập dữ liệu qua API.

Phương pháp 2: Thu thập dữ liệu theo trang

Phương pháp này cho phép thu thập dữ liệu theo trang (sites), trong đó các trang có dữ liệu muốn thu thập phải được liệt kê trong hệ thống. Hình 4 và hình 5 mô tả các trang được liệt kê trong hệ thống của chúng tôi để thử nghiệm phương pháp thu thập này. Phương pháp này giúp thu thập toàn bộ dữ liệu của các trang được liệt kê.



Hình 5. Một số sites được liệt kê tại Thái Nguyên

Phương pháp thu thập dữ liệu theo trang được thực hiện bằng 2 cách: Liệt kê danh sách trang (panel) và Tự động lan tỏa.

Cách 1: Thu thập theo phương pháp liệt kê danh sách trang [8]

Từ danh sách các trang, nhóm chịu trách nhiệm về dữ liệu của chúng tôi sẽ cài đặt, thiết lập các con nhện (crawl) để quét các trang liên tục và thực hiện sao chép các thảo luận của người dùng trên trang đó. Crawl thao tác như một con người, sau khi truy cập vào trang cần thu thập nó sẽ tự động quét nội dung của trang, nhận diện đâu là bài viết, sao chép toàn bộ nội dung được hiển thị trong bài viết như: Title, Replies, Authors, URL,... Crawl sẽ quay lại mỗi trang dựa vào thời gian được thiết lập để tìm ra bài viết mới và các replies mới cho bài viết cũ.

Crawl chỉ có thể thấy được những thông tin ở chế độ công khai, thu thập được những thảo luận để chế độ public, chứ không lấy được các thảo luận private, tuân thủ theo luật privacy. Tuy nhiên, crawl có thể lấy được thảo luận trong nhóm Facebook kín, bằng cách đăng nhập bằng một “member ID” (**Bot**) của group đó, nhưng việc này cần có sự đồng ý của quản trị viên của nhóm đó.

Bot [5] là một chương trình tự động tương tác với người dùng và được tích hợp trên fanpage. Bot cơ bản gồm có 2 loại: Thứ nhất là Bot dựa trên kịch bản. Đây là loại Bot đơn giản nhất và cũng dễ dùng nhất, chỉ cần tạo

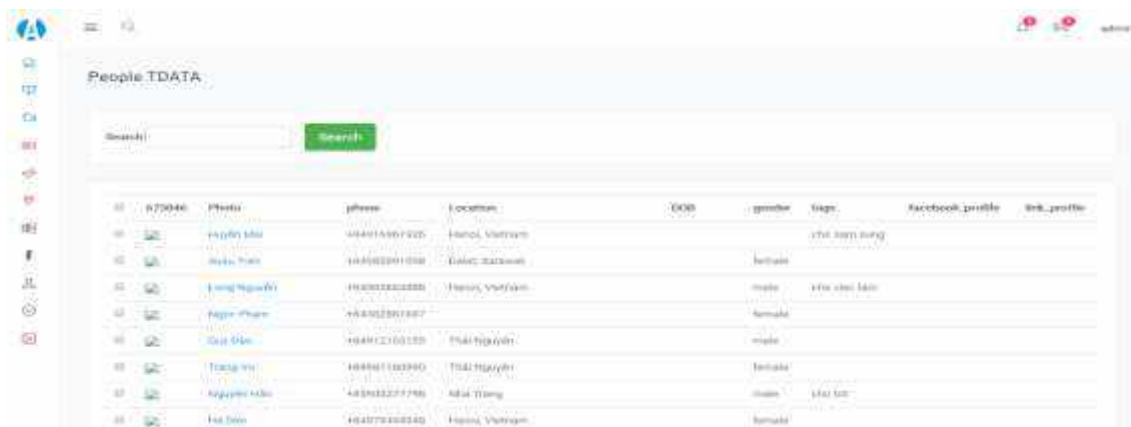
trước những kịch bản để Bot dựa vào đó tương tác với người dùng. Loại này rất hiệu quả trong việc hỗ trợ đặt hàng, thu thập thông tin phản hồi. Loại thứ 2 là Bot dựa trên trí tuệ nhân tạo. Loại này hiểu ngôn ngữ của con người và học hỏi khi trò chuyện với người dùng, đây là một lĩnh vực thuộc máy học (machine learning). Loại này rất hữu dụng trong việc chăm sóc khách hàng.

Mục đích của Bot giúp tiết kiệm sức người, chi phí thuê nhân viên. Các Bot sẽ theo dõi các trang của người dùng, sau đó thu thập thông tin từ người dùng bao gồm: Hồ sơ cá nhân như tên, tuổi, giới tính, nghề nghiệp, nơi làm việc...; sở thích; hành vi; nhu cầu... Trong thời gian thử nghiệm chúng tôi đã thu thập được thông tin của 673046 người dùng thông qua mạng xã hội facebook (hình 6).

Đối với các Bot cần thiết lập các kịch bản cho Bot, cụ thể chính là chức năng Schedule, và duy trì sự sống của Bot (hình 7).

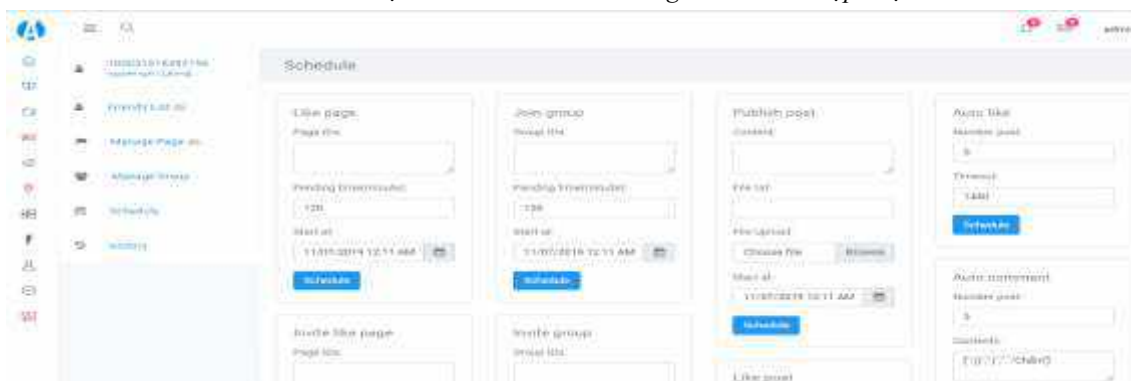
Ưu điểm của phương pháp này là có thể thu thập được 100% dữ liệu trong trang. Tuy nhiên phương pháp này cũng có nhiều nhược điểm:

- Phụ thuộc vào đường truyền internet: Tốc độ trả dữ liệu của trang, khả năng nhận diện nội dung và cấu trúc trang của crawl.
- Cấu trúc site thay đổi: Các trang diễn đàn lớn thường có sự thay đổi về cấu trúc hàng năm nên khi crawl gặp cấu trúc mới khác với thiết kế ban đầu thì sẽ dẫn đến việc thu thập bị gián đoạn.



ID	Photo	Name	Phone	Location	DOB	Gender	Age	Facebook profile	Link profile
673046		Nguyễn Thị	04411567325	Hanoi, Vietnam				chị bạn cùng	
		Nguyễn Thị	04402201038	Hanoi, Vietnam		female			
		Nguyễn Thị	04402201038	Hanoi, Vietnam		male	chị bạn cùng		
		Nguyễn Thị	04402201037	Hanoi, Vietnam		female			
		Nguyễn Thị	04402201038	Hanoi, Vietnam		male			
		Nguyễn Thị	04402201038	Hanoi, Vietnam		female			
		Nguyễn Thị	04402201038	Hanoi, Vietnam		male			
		Nguyễn Thị	04402201038	Hanoi, Vietnam		female			

Hình 6. Dữ liệu cá nhân của khách hàng mà Bot thu thập được



Hình 7. Thiết lập các chức năng cho Bot

- Quản trị viên của các sites thiết lập cơ chế nhận diện và chặn việc thu thập dữ liệu nếu việc thu thập dữ liệu tự động lạm dụng tài nguyên của các website và gây ảnh hưởng đến băng thông. Các Crawl cũng thường xuyên phải cập nhật và thay đổi tính danh để vượt qua các cơ chế chặn này.

- Không thể truy cập để thu thập dữ liệu khi trang đang bảo trì hoặc xảy ra vấn đề.

Vì những lý do trên, việc thiếu dữ liệu hay gián đoạn trong quá trình thu thập dữ liệu là điều không thể tránh khỏi nên chúng tôi có một đội ngũ lập trình viên (data team) làm việc liên tục để cập nhật Crawl, xử lý các tình huống bất thường xảy ra để đảm bảo đầy đủ dữ liệu, đặc biệt trong các trường hợp doanh nghiệp chạy chiến dịch marketing hay xử lý khi có khủng hoảng truyền thông.

Đây cũng là lý do các công cụ thu thập dữ liệu của nước ngoài hay các công cụ miễn phí không thể hoạt động hiệu quả ở Việt Nam do thường xuyên bị thiếu dữ liệu và việc thu thập

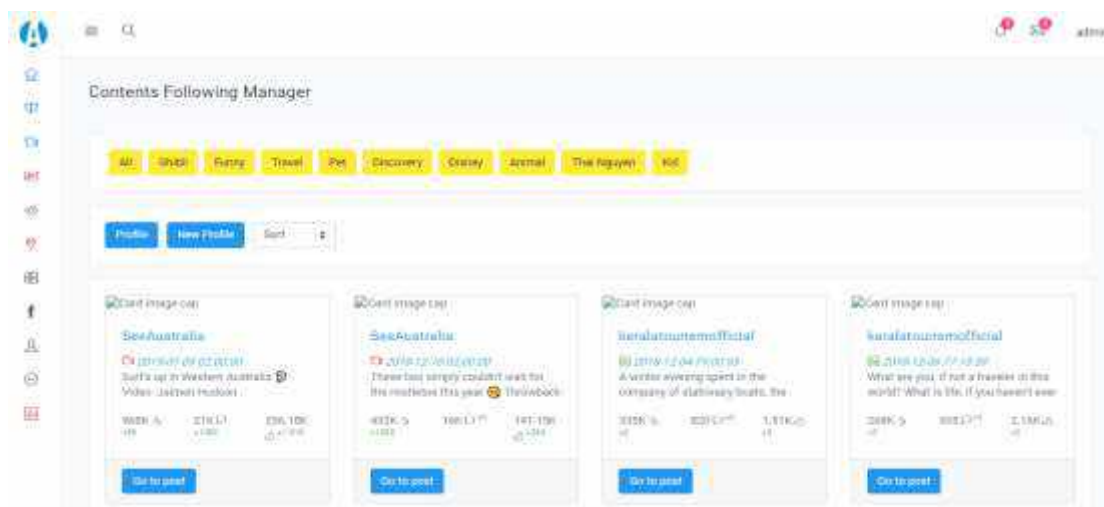
danh sách sites không đủ, hoặc khi gặp một trong các vấn đề trên thì không có nhân sự để giải quyết kịp thời.

Cách 2: Thu thập theo phương pháp trang tự động lan tỏa

Việc thu thập trang tự động có thể được thực hiện bằng 2 cơ chế thông minh:

- Thu thập theo trend: Khi có một bài viết, sự kiện, chủ đề, xu hướng được nhắc đến nhiều nhất trên mạng xã hội, hệ thống sẽ tự động phát hiện và thu thập các trang có chứa thảo luận về chủ đề đó. Ví dụ, khi có một sự kiện được báo chí đưa tin nhiều, hệ thống sẽ tự động phát hiện từ khóa về sự kiện và thu thập thông tin xung quanh sự kiện đó. Đối với mỗi bài viết, dữ liệu mà chúng tôi thu nhận được là thông tin về bài viết, số lượt like, comment, chia sẻ bài viết đó (hình 8).

- Thu thập theo cơ chế lan tỏa: Từ những trang/group đã thu thập được, hệ thống sẽ phát hiện và thu thập các trang/group/user khác được nhắc đến trong những trang này.



Hình 8. Kết quả thu thập nội dung/ trang có chứa nhiều thảo luận

Thu thập dữ liệu theo trend và theo cơ chế lan tỏa là 2 quá trình được thực hiện đồng thời, để đảm bảo có thể thu thập được nhanh nhất và đầy đủ nhất dữ liệu về các chủ đề đang được quan tâm thảo luận nhiều trên mạng xã hội.

4. Kết luận

Đối với các doanh nghiệp thì điều họ quan tâm nhất là khách hàng của họ nghĩ gì về sản phẩm. Khách hàng thời đại công nghệ 4.0 này họ rất thông minh, mẫn cảm với tin nhắn rác và các cuộc trò chuyện, comment ảo trong bài post, các quảng cáo không tự nhiên. Khách hàng chỉ tiếp nhận thông tin thực sự liên quan đến họ, đáp ứng đúng nhu cầu của họ tại đúng thời điểm. Bằng việc áp dụng sự phát triển của công nghệ thông tin, công nghệ số và các công nghệ tiên tiến trong thời đại công nghiệp 4.0. Chúng tôi đã tạo ra những nhân viên vô hình có thể thay thế con người thực hiện một số thao tác đơn giản, lặp đi lặp lại nhiều lần. Hơn nữa khi áp dụng tools này để thu thập dữ liệu các doanh nghiệp vừa và nhỏ ở Việt Nam đã tiết kiệm được rất nhiều chi phí cũng như thời gian, từ đó có thể bắt kịp được với xu thế phát triển nhanh chóng của công nghệ.

Lời cảm ơn

Bài báo là sản phẩm của đề tài có mã số T2019-07-08, được tài trợ bởi kinh phí của trường Đại học Công nghệ Thông tin và Truyền thông – ĐHTN. Nhóm tác giả xin chân thành cảm ơn sự tài trợ của quý Nhà trường.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1]. M. Johnsen, *The future of artificial intelligence in digital marketing*. United States of America, ISBN:978-1976001062, 2017.
- [2]. Simon Kemp, "Digital in 2018: world's internet users pass the 4 billion mark", *We Are Social Ltd*, 2018. [Online]. Available: <https://wearesocial.com/blog/2018/01/global-digital-report-2018>. [Accessed Jan. 30, 2020].
- [3]. M. H. Pattankar, *Social Media Marketing*. Lulu publication, ISBN:978-1-387-71111-6, 2018.
- [4]. Buzzmetrics Joint Stock Company, "Insight the target customers: How social listening tools collect data?", *Buzzmetrics Joint Stock Company* 2014. [Online]. Available: <https://buzzmetrics.com/>. [Accessed Jan. 30, 2020].
- [5]. James Choi, "The 29 Best AI Chatbot Building Platforms For Any Business", *chatbots.org*. [Online]. Available: <https://chatbots.org/best-ai-chatbot-tools#> [Accessed Jan. 30, 2020].
- [6]. S. Gregory, *Finding Overlapping Communities Using Disjoint Community Detection Algorithms*. CompleNet, pp. 47-61, 2009.
- [7]. C. Wang, J. Han, and R. Discovery, *Encyclopedia of Social Network Analysis and Mining*. Springer, pp. 1589-1598, 2014.
- [8]. Q. T. Ha, and M. V. Tran, *Analysis and mining social network*, VNU - University of Engineering and Technology, 2016.